2016

Mechanisms of Extracellular Oncogenic Dysregulation and Antibody Targeting of the Epidermal Growth Factor Receptor

Atrish Bagchi
University of Pennsylvania, atrish.bagchi@gmail.com

Mechanisms of Extracellular Oncogenic Dysregulation and Antibody Targeting of the Epidermal Growth Factor Receptor

Abstract
Regulation of the Epidermal Growth Factor Receptor (EGFR) by its growth factor ligands is critical in many biological processes, including development and tissue maintenance and growth. Aberrant overexpression or activation by mutation of EGFR is associated with many human tumors. In these contexts, constitutive signaling can lead to cellular transformation and oncogenesis, thereby driving the cancer. The EGFR is the target of several existing or developing cancer therapies or immunotherapies, including monoclonal antibodies that prevent its activation. Activating mutations in cytoplasmic tyrosine kinase domain have been identified in many cancers, and have been the focus of mechanistic work. In this dissertation, I focus on the mode of oncogenic dysregulation by novel extracellular mechanisms.

Extracellular oncogenic variants of EGFR include point mutations and alternative splice variants of EGFR. I find through biochemical analysis of the activating missense mutations in the extracellular region of EGFR that the soluble extracelluar region of EGFR (sEGFR) harboring these mutations bind ligands with elevated affinities. The dimerization energetics of these sEGFR mutants is not measurably altered, which suggests that additional interactions from the membrane and/or the intracellular region are important to this novel mode of extracellular oncogenic dysregulation of EGFR. I present preliminary progress towards the application of hydrogen/deuterium exchange coupled to mass spectrometry to analyze such allosteric (dys)regulation of the EGFR.

In a second focus, I studied mechanisms of antibody targeting of EGFR. There are several monoclonal therapeutic antibodies that are in clinical development or use that target the EGFR/ErbB/HER family of receptor tyrosine kinases, including cetuximab/Erbitux™, panitumumab/Vectibix™, and necitumumab/Portrazza™, which all target EGFR, as well as trastuzumab/Herceptin™ and pertuzumab/Perjeta™, which target ErbB2/HER2. Unfortunately, as observed for most targeted therapies for cancer, resistance to these antibody therapies limits the duration of their effective treatment. Recent exome sequencing analyses of KRAS-WT colorectal cancer patients resistant to cetuximab treatment has identified epitope mutations as a mechanism of resistance. Whereas these mutated receptors bind cetuximab with dramatically decreased affinities, I report that they retain high affinity binding for necitumumab, a humanized IgG1 anti-EGFR antibody that shares the same epitope as cetuximab and panitumumab, and was recently FDA approved for squamous non-small cell lung carcinoma. I determined an X-ray crystal structure of the Fab fragment of necitumumab with the most commonly found resistance mutation—S492R (or S468R using the numbering scheme that starts at the beginning of the mature EGFR protein). This structure reveals a relatively hydrophobic cavity in the paratope of necitumumab that can accommodate the arginine at position S492/468 in the EGFR epitope. Further I find that other cetuximab and panitumumab resistance variants of EGFR are also permissive for necitumumab binding, suggesting significant plasticity in binding of necitumumab to EGFR. A survey of structures of therapeutic antibodies bound to their targets suggests that paratope shape may be an important property to consider in the selection of monoclonal antibodies in therapeutic strategies.

Another mechanism of oncogenic dysregulation is the gene rearrangement of EGFR that results in EGFR variant III (EGFRvIII), an important target of many classes of immunotherapies for glioblastoma multiforme (GBM). I show in small angle X-ray scattering analyses of the ectodomain of EGFRvIII some evidence of structural flexibility in domain II that may be important for its documented transactivation of other receptor tyrosine kinases. I also report an X-ray crystal structure of the ectodomain of EGFRvIII in complex with the

antigen binding or VHH domain of a camelid heavy-chain only antibody (HCAb), that has ~25-fold specificity for EGFRvIII compared to wild type EGFR. The structure reveals that the VHH gains specificity for EGFRvIII by targeting an epitope on domain IV that is sterically occluded in wild type EGFR by the intramolecular 'tether'. This structure provides the direct evidence of dynamic uncoupling of the 'tether'. My work corroborates the utility of the 'tether' as a source of antibody specificity for oncogenic EGFR, and is the first structural view of specific antibody targeting of an oncogenic EGFR variant.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Biochemistry & Molecular Biophysics

First Advisor
Kathryn M. Ferguson

Second Advisor
Gregory D. Van Duyne

Keywords
antibody, cancer, drug, EGFR, oncogenic mutation, resistance

Subject Categories
Biochemistry | Biophysics

MECHANISMS OF EXTRACELLULAR ONCOGENIC DYSREGULATION AND ANTIBODY TARGETING OF THE EPIDERMAL GROWTH FACTOR RECEPTOR

Atrish Bagchi

A DISSERTATION

in

Biochemistry and Molecular Biophysics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2016

Supervisor of Dissertation

_____

Kathryn M. Ferguson

Adjunct Associate Professor of Physiology

Co-Supervisor

_____

Gregory D. Van Duyne

Professor of Biochemistry & Biophysics

Graduate Group Chairperson

_____

Kim A. Sharp, Associate Professor of Biochemistry & Biophysics

Dissertation Committee

Ben E. Black, Associate Professor of Biochemistry and Biophysics

Benjamin A. Garcia, Presidential Associate Professor of Biochemistry and Biophysics

Mark I. Greene, John Eckman Professor of Medical Science and Professor of Pathology and Laboratory Medicine

Ronen Marmorstein, Professor of Biochemistry and Biophysics

James Shorter, Associate Professor of Biochemistry and Biophysics

Zhihong Wang, Assistant Professor of Chemistry and Biochemistry, University of the Sciences in Philadelphia

MECHANISMS OF EXTRACELLULAR ONCOGENIC DYSREGULATION AND ANTIBODY TARGETING OF THE EPIDERMAL GROWTH FACTOR RECEPTOR

DEDICATION

To my parents, Sumita and Harish, who sparked and fostered my interest in scientific inquiry, and always challenged me to be more curious.

ACKNOWLEDGMENT

I thank my mentor Kate, and all of my labmates in the Ferguson and Lemmon laboratories for making the past few years a very fun journey and rewarding experience, and always challenging me to think of and perform a better experiment. I must say that I would not trade my experience in this lab for any other, as I have grown and learned more than I ever could have possible imagined at the beginning. I especially thank Nick Bessman, with whom I collaborated in work presented in Chapter Two, and who taught me many techniques, and was always available to help and answer questions when I was new in the lab. I thank many collegial interactions with the Lemmon lab, especially Neo Wu, with whom I had many productive interactions and discussions with regards to HDX/MS analysis. I also thank Jeannine, Camilla, Kelsey, Dan, Jin, Yu-San, and Pam for many fun late night conversations, and advice (both experimental and career). I also thank Steve Stayrook for help with SAXS and home source X-ray diffraction analyses. I am very grateful to Jason Moore and all of the beamline staff at Argonne National Labs for guidance and training for X-ray diffraction experiments. I also very gratefully thank Greg Van Duyne and Kushol Gupta for experimental and career advice, and for generously adopting me as a student when my lab moved to Yale. I thank the entire Physiology and Crystal Talk communities, especially the members of the Black, Marmorstein, and Dominguez labs, for many fun and stimulating collegial interactions. I also thank all of my colleagues in the Englander lab, especially Zhongyuan Kan, for teaching me how to do HDX and all of the invaluable experimental advice on the HDX project. I thank Pam Burgess-Jones, and all of the staff in the BMB graduate group, especially Ruth Keris, for crucial administrative support. I thank Josh Wand for support through the Structural Biology and Molecular Biophysics training grant, which funded me for two years, and for coordinating a very diverse and interesting group of speakers for the SBTG lunches. I also thank many BMB graduate group faculty, especially my

former lab neighbors Kenji Murakami and Jeremy Wilusz, as well as Kristen Lynch, for many stimulating collegial interactions and career advice. I am also grateful for informal discussions with Roland Dunbrack about antibody structures. I also thank all of my closest friends in graduate school – there are too many to individually name, but I especially give a shout out to my classmates Cory, Francesca, Chris, Cat, John, Jarrett, Zhenyu, and Glennis. Thank you so much for being there for me, attending practice talks, reading grants and papers, giving critical feedback, and most importantly being great friends. I thank all of the wonderful baristas at Greenstreet Coffee Roasters at 11th and Spruce, where I basically lived for the past few months while writing my thesis. I thank all of my friends in Philadelphia and beyond for personal and intellectual support, and for making life very fun, even in the distressful times I have had in the past year. I can very honestly say that I would not be the person or scientist I am if I had not come here, to the city of Philadelphia, and I reflect on my decision to come here very positively. Last but not least, I thank my parents for their never-ending love and support, and keeping me sane through the past few years, especially in the last year. Thank you!

ABSTRACT

MECHANISMS OF EXTRACELLULAR ONCOGENIC DYSREGULATION AND ANTIBODY TARGETING OF THE EPIDERMAL GROWTH FACTOR RECEPTOR

Atrish Bagchi

Kathryn M. Ferguson

Regulation of the Epidermal Growth Factor Receptor (EGFR) by its growth factor ligands is critical in many biological processes, including development and tissue maintenance and growth. Aberrant overexpression or activation by mutation of EGFR is associated with many human tumors. In these contexts, constitutive signaling can lead to cellular transformation and oncogenesis, thereby driving the cancer. The EGFR is the target of several existing or developing cancer therapies or immunotherapies, including monoclonal antibodies that prevent its activation. Activating mutations in cytoplasmic tyrosine kinase domain have been identified in many cancers, and have been the focus of mechanistic work. In this dissertation, I focus on the mode of oncogenic dysregulation by novel extracellular mechanisms.

Extracellular oncogenic variants of EGFR include point mutations and alternative splice variants of EGFR. I find through biochemical analysis of the activating missense mutations in the extracellular region of EGFR that the soluble extracelluar region of EGFR (sEGFR) harboring these mutations bind ligands with elevated affinities. The dimerization energetics of these sEGFR mutants is not measurably altered, which suggests that additional interactions from the membrane and/or the intracellular region are important to this novel mode of extracellular oncogenic dysregulation of EGFR. I present preliminary progress towards the application of hydrogen/deuterium exchange coupled to mass spectrometry to analyze such allosteric (dys)regulation of the EGFR.

In a second focus, I studied mechanisms of antibody targeting of EGFR. There are several monoclonal therapeutic antibodies that are in clinical development or use that target the EGFR/ErbB/HER family of receptor tyrosine kinases, including cetuximab/Erbitux®, panitumumab/Vectibix®, and necitumumab/Portrazza®, which all target EGFR, as well as trastuzumab/Herceptin® and pertuzumab/Perjeta®, which target ErbB2/HER2. Unfortunately, as observed for most targeted therapies for cancer, resistance to these antibody therapies limits the duration of their effective treatment. Recent exome sequencing analyses of KRAS-WT colorectal cancer patients resistant to cetuximab treatment has identified epitope mutations as a mechanism of resistance. Whereas these mutated receptors bind cetuximab with dramatically decreased affinities, I report that they retain high affinity binding for necitumumab, a humanized IgG1 anti-EGFR antibody that shares the same epitope as cetuximab and panitumumab, and was recently FDA approved for squamous non-small cell lung carcinoma. I determined an X-ray crystal structure of the Fab fragment of necitumumab with the most commonly found resistance mutation—S492R (or S468R using the numbering scheme that starts at the beginning of the mature EGFR protein). This structure reveals a relatively hydrophobic cavity in the paratope of necitumumab that can accommodate the arginine at position S492/468 in the EGFR epitope. Further I find that other cetuximab and panitumumab resistance variants of EGFR are also permissive for necitumumab binding, suggesting significant plasticity in binding of necitumumab to EGFR. A survey of structures of therapeutic antibodies bound to their targets suggests paratope shape may be an important property to consider in the selection of monoclonal antibodies in therapeutic strategies.

Another mechanism of oncogenic dysregulation is the gene rearrangement of EGFR that results in EGFR variant III (EGFRvIII), an important target of many classes of immunotherapies for glioblastoma multiforme (GBM). I show in small angle X-ray scattering analyses of the

ectodomain of EGFRvIII some evidence of structural flexibility in domain II that may be important for its documented transactivation of other receptor tyrosine kinases. I also report an X-ray crystal structure of the ectodomain of EGFRvIII in complex with the antigen binding or VHH domain of a camelid heavy-chain only antibody (HCAb), that has ~25-fold specificity for EGFRvIII compared to wild type EGFR. The structure reveals that the VHH gains specificity for EGFRvIII by targeting an epitope on domain IV that is sterically occluded in wild type EGFR by the intramolecular 'tether'. This structure provides the direct evidence of dynamic uncoupling of the 'tether'. My work corroborates the utility of the 'tether' as a source of antibody specificity for oncogenic EGFR, and is the first structural view of specific antibody targeting of an oncogenic EGFR variant.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER ONE

Introduction

A longstanding question in signal transduction biology is how cell surface receptors propagate the message of ligand availability across the plasma membrane. Regulation of growth factor receptors by their growth factor ligands is critical in control of cellular proliferation, tissue development, and angiogenesis, and contributes to many developmental processes (Cote, 2012; Maina, 1996; McDonell, 2015). An important class of growth factor receptors contain intracellular tyrosine kinase domains, which function to mediate trans-autophosphorylation of regulatory tyrosines (Lemmon and Schlessinger, 2010). Whereas the catalytic kinase domains are structurally and functionally well conserved, these receptor tyrosine kinases (RTK's) contain diverse extracellular domain architectures, with many combinations of different types of subdomains. The structural diversity of the extracellular architectures of receptor tyrosine kinases suggests mechanistic diversity in how these receptors are regulated by their cognate growth factor ligands, and such diversity is now starting to be fully explored (Lemmon and Schlessinger, 2010; Lemmon, 2014).

The epidermal growth factor receptor (EGFR) was the first receptor tyrosine kinase to be identified (Carpenter, 1978) and mechanistically investigated. It was also the first growth factor receptor for which the concept of ligand-induced dimerization was shown as a relevant mechanism for extracellular regulation of its intracellular tyrosine kinase domain (Lemmon, 1997; Yarden, 1987). Epidermal growth factor (EGF) induces receptor-mediated dimerization of EGFR, which leads to activation of its cytoplasmic tyrosine kinase domains. Subsequent autophosphorylation of regulatory tyrosines in a disordered cytoplasmic C-terminal regulatory region results in recruitment of adaptor proteins that contain Src homology 2 (SH2) and phosphotyrosine binding (PTB) domains. This recruitment initiates cascades of downstream signaling events through multiple pathways, including MAP kinase/Erk, phosphatidylinositol 3-kinase/Akt, and phospholipase C/$Ca^{2+,}$ culminating in a variety of cellular responses, including proliferation or differentiation (Lemmon and Schlessinger, 2010).

Ligand-induced activation of EGFR can be bypassed by genetic mutation or its overexpression. The resulting elevated aberrant activity of the EGFR can lead to constitutive activation of proliferative signaling that causes oncogenic transformation. Moreover, overexpression and mutation of EGFR have been implicated in and drive many human cancers, including tumors of the brain, head and neck, and epithelial tissues. The most well characterized, and common, of EGFR-activating oncogenic mutations are found in its tyrosine kinase domain, and include hydrophobic residues in a helix that stabilize an activation loop that sterically locks the catalytic site in its inactive state. This activating mechanism drives EGFR's responsiveness to tyrosine kinase inhibitor gefitinib (Lynch, 2004; Paez, 2004). The mechanistic basis for the mode of EGFR activation by these mutations is reasonably well established for many of the mutations that have been identified (Shan, 2012), though gaining predictive power of mutations that are activating and driver mutations remains an important avenue of investigation. This dissertation focuses on structural mechanisms of extracellular oncogenic dysregulation and therapeutic antibody targeting of the epidermal growth factor receptor family.

The structural basis for ligand-induced dimerization of the epidermal growth factor receptor family was pioneered by X-ray crystallographic studies in the early 2000's, which revealed at atomic resolution how EGF binding to its receptor leads to receptor-mediated dimerization of its extracellular region (Garrett, 2002; Ogiso, 2002). Ligand binding to EGFR induces dramatic conformational rearrangement in its extracellular region in which an intramolecular 'tether' between cysteine rich domains (CRD's) II and IV is broken, bringing domains I and III closer together and exposing a critical β hairpin in cysteine rich domain II that mediates important dimerization contacts (Burgess, 2003; Ferguson, 2003). These X-ray crystal structures provided a structural model for ligand-induced homodimerization of EGFR (Figure 1.1) (Burgess, 2003) and a framework for study of relationships between ligand binding to ErbB receptors and receptor homo- or hetero-oligomerization (Ferguson, 2000;

Kumagai et al., 2001). Structural studies of other members of the ErbB/HER family showed that similar intramolecular 'tethers' between domains II and IV also exist in ErbB3 (Cho, 2002), and ErbB4 (Bouyain, 2005). By contrast, the orphan receptor ErbB2/HER2, for which there is no known ligand, was found by crystallographic analysis to be constitutively 'extended' as a monomer, revealing that it is structurally distinct from the other ErbB receptors (Cho, 2003). This structural distinction can be harnessed by targeted antibody therapeutics to gain specificity for targeting ErbB2/HER2, such as trastuzumab/Herceptin™ (Cho, 2003) and pertuzumab/Perjeta™ (Franklin, 2004). In Chapter Five, I discuss antibody targeting of an oncogenic EGFR variant that utilized a similar binding mechanism.

Hypotheses regarding the precise biological role of this intramolecular 'tether' in ErbB receptors have evolved since these first structures were determined. Though in 2003, it was proposed that the domain II/IV intramolecular interaction may be responsible for receptor autoinhibition (Burgess, 2003; Ferguson, 2003; Walker, 2004), mutation of tethering residues is not activating (Mattoon, 2004) and the extracellular regions of both EGFR and ErbB4 retain a tethered-like conformation in the absence of these interactions (Dawson, 2007; Liu, 2012). Simple exposure of the β-hairpin in domain II by deletion of domain IV is not sufficient to drive ligand-independent dimerization of the EGFR extracellular region, suggesting that ligand-induced conformational remodeling of domain II is critical for EGFR activation (Dawson, 2005). In addition, the unliganded, inactive *Drosophila melanogaster* EGFR extracellular region lacks a domain II/IV tether and is instead an 'extended' monomer, suggesting that the presence of a 'tether' is not a sufficient autoinhibitory mechanism (Alvarado, 2009).

In 2006, Mellinghoff and colleagues identified activating missense mutations in the extracellular region of the EGFR in glioblastoma multiforme (GBM) tumors or cell lines. Expression of these altered EGFR molecules induces cellular transformation of NIH3T3 cells and tumor growth in murine xenograft models (Lee et al., 2006). In addition, these mutations are sufficient to induce IL-3 independent growth of the murine B cell line Ba/F3, which are

dependent on IL-3 for growth unless they express a constitutively active kinase (Warmuth, 2007). These mutations are largely located in critical 'hinge' points (Lee et al., 2006; Ymer, 2011); clustering in a hydrophobic interface between domains I and II or at the interface between domains II and III – regions that had previously been ignored (Figure 1.2). Subsequent exome sequencing analysis of a panel of other epithelial cancers identified analogous mutations in ErbB3 in the same interface, suggesting possible mechanistic similarities in the modes of extracellular oncogenic dysregulation of ErbB receptors (Jaiswal et al., 2013). In work published in 2014 (Bessman, 2014) and reproduced in Chapter Two, I found that these EGFR-activating mutations in the extracellular region of EGFR increase ligand binding affinities for the mutated EGF receptors without measurably altering the ligand-induced EGFR homo-dimerization affinities. As described in Bessman et al. (2014) and Chapter Two, these data suggest that the extracellular oncogenic mutations may alter the thermodynamic linkage of ligand binding and dimerization of EGFR rather than simply promoting ligand-independent dimerization (Bessman, 2014).

Structural studies of the intracellular EGFR tyrosine kinase domain showed that, unlike some other tyrosine kinases that require phosphorylation of an activation loop to stabilize an active conformation (Hubbard, 2013; Lemmon, 2014), formation of an asymmetric kinase dimer is essential for EGFR tyrosine kinase activation (Figure 1.1). In this asymmetric dimer, the C-lobe of one, inactive, monomer interacts with the N-lobe of the other and thereby allosterically stabilizes its catalytically competent conformation (Zhang, 2006). The formation of this asymmetric dimer is required for tyrosine kinase activation as mutation of the crystallographic dimer interface abolishes kinase activity. It is also clear that formation of this asymmetric dimer can be regulated by mechanisms involving interactions with the juxtamembrane regions of EGFR. It has been shown through crystallographic analysis that the intracellular juxtamembrane region forms a 'latch' around this asymmetric dimer, stabilizing it through van der Waals interactions (Figure 1.1) (Jura, 2009; Red Brewer, 2009).

Despite these advances in our structural knowledge of isolated EGFR domains, the precise nature of the allosteric coupling of extracellular ligand-induced dimerization and intracellular dimerization of the tyrosine kinase domains remains unclear. The extracellular region of EGFR was originally viewed as a module that served simply to drive ligand-induced dimerization of the intracellular tyrosine kinase domain. However, it has been shown in recent years that inactive, EGFR dimers also exist that are altered by ligand binding (Chung, 2010; Yu, 2002). It is also unclear from structural studies how structural asymmetry could be communicated across the plasma membrane, as structures of ligated human EGFR in complex with EGF or TGFα are completely symmetric (Lu et al., 2012; Ogiso, 2002). Ligand binding studies of intact EGFR on the cell surface argue that negative cooperativity of ligand binding underlies this coupling of asymmetry (MacDonald, 2008; Pike, 2012). The implicit structural asymmetry associated with negative cooperativity was crystallographically observed in a ligated structure of the Spitz-bound EGFR ortholog from *Drosophila melanogaster* (Alvarado et al., 2010). However, in the human counterpart, such negatively cooperative ligand binding is not recapitulated in studies of the isolated EGFR extracellular region, suggesting that additional interactions from the cytoplasmic domains and/or the plasma membrane are also important for such complex ligand binding equilibria. In addition, the extracellular regions of the *Drosophila melanogaster* and *C. elegans* orthologs of EGFR dimerize with high affinity in the absence of ligand (Freed, 2015), observations that suggest there may be autoinhibitory mechanisms within receptor dimers that are altered by ligand binding.

Taken together, it is clear that EGFR is allosterically regulated in a manner similar to other allosterically regulated receptor tyrosine kinases, such as the insulin receptor (Lemmon, 2014). In Chapter Three, I describe advances towards application of structural techniques such as hydrogen/deuterium exchange coupled to mass spectrometry to understand this

allostery. Future work that will build upon this may reveal how extracellular oncogenic mutations in EGFR dysregulate the coupling of extra- and intra-cellular communication.

Another mode of extracellular oncogenic dysregulation of EGFR results from gene rearrangements of EGFR and expression of exon excluded forms of the receptors. Historically, one of the first oncogenes to be identified was v-ErbB, from avian erythroblastosis virus (from which the ErbB family derived its name). v-ErbB is a mutated form of EGFR that results in a large deletion of its extracellular region and also harbors several intracellular region mutations. The most common extracellular oncogenic form of EGFR produces a similar, N-terminally truncated form of the receptor, called variant III (EGFRvIII) (Sugawa, 1990; Wong, 1992). EGFRvIII is constitutively active and phosphorylated and has been found to contribute cellular transformation and therapeutic resistance in glioblastoma (Gan, 2009), non-small cell lung cancer (Okamoto, 2003) and head and neck squamous cell carcinoma (Sok, 2006). Because of its expression at high frequency in glioblastoma, EGFRvIII has emerged as an important target for several classes of cancer immunotherapies (Choi, 2013; Johnson, 2015). In Chapter Five, I identify a novel mode of specific targeting of EGFRvIII by a nanobody that may have potential as an imaging and/or therapeutic agent.

Due to the important role of the epidermal growth factor receptor family in driving human cancers, the members of this family are the targets of several classes of cancer therapies. These include therapeutic antibodies, such cetuximab that block ligand-induced activation of EGFR (Li, 2008; Li, 2005; Schmeidel, 2008; Schmitz, 2009), and small molecule tyrosine kinase inhibitors that block ATP binding and kinase activation (Gazdar, 2009). In the context of both targeted therapeutic approaches, clinical resistance remains a formidable problem. Resistance can result from gene amplification, missense mutations, or alteration of downstream signaling networks such that they are rendered less sensitive to the input signals from growth factor receptors.

In the case of the EGFR, there are three targeted anti-EGFR antibody therapies currently approved by the U.S. Food and Drug Administration for clinical use in cancer therapy. These antibodies are cetuximab/Erbitux™, panitumumab/Vectibix™, and necitumumab/Portrazza™. Cetuximab/Erbitux, a mouse/human chimeric IgG1 antibody approved for head and neck and colorectal cancers, binds to EGFR domain III to an epitope that overlaps the binding site for EGF and other EGFR ligands (Li, 2005). Cetuximab blocks ligand binding to EGF receptor. It further sterically prevents the ligand-induced conformational changes required for EGFR dimerization. Panitumumab is a humanized IgG2 antibody that is also approved for head and neck and colorectal cancers. The efficacies of these antibody therapies for treating metastatic colorectal cancers was recently compared in a phase III clinical trial called ASPECCT (Price, 2015). Necitumumab, a humanized IgG1 antibody that shares the same framework region as cetuximab, binds to a nearly identical epitope as cetuximab, though through completely different binding interactions with its paratope (Li, 2008). Necitumumab/Portrazza™ in combination with gemcitabine and cisplatin for treatment of squamous non-small cell lung cancer received approval by the U.S. Food and Drug Administration in November 2015.

Recent sequencing efforts of patients resistant to EGFR-directed antibody therapies has revealed that one of the major mechanisms of resistance to these targeted therapies are point mutations in the shared epitope of these three antibodies (Bertotti, 2015). These missense mutations decrease the effectiveness of these antibody inhibitors of EGFR by substantially decreasing the binding affinities of the antibodies for the mutated receptor molecules. Sixteen percent of patients in the cetuximab arm of the ASPECCT trial developed resistance through one particular mutation – S492R in pro-EGFR, or S468R using the mature amino acid numbering of EGFR (Bertotti, 2015; Montagut, 2012; Price, 2015). In Chapter Four, I present a collaborative study with Eli Lilly & Company to investigate the binding of necitumumab to EGFR harboring cetuximab resistance mutations. My structural analysis suggests that paratope shape may be an important feature that dictates the susceptibility of a

targeted antibodies to development of resistance due to epitope mutation. Surprisingly, the interfaces of cetuximab and necitumumab in complex with the EGFR are not significantly different in quantitative measures of shape complementarity, despite the latter containing a buried hydrophobic cavity that contributes to its structural plasticity (Lawrence, 1993). Comparison of the paratopes of cetuximab and necitumumab to those of other therapeutic antibodies for which complex structures have been determined reveals a structural class of therapeutic antibodies that contain similar buried cavities between their VL and VH domains. Further studies are required to establish whether antibodies this class of antibody paratope are less susceptible to development of resistance due to epitope mutations. This work has implications for therapeutic antibody design, and the clinical use of necitumumab for these cases of cetuximab resistance.

I focus on aspects of two areas in this dissertation: (1) the structural and mechanistic basis for extracellular oncogenic dysregulation of the epidermal growth factor receptor (EGFR), and (2) the structural basis of antibody targeting and/or inhibition of the EGFR, either in the context of oncogenic dysregulation, or resistance to existing monoclonal antibody therapies. My studies highlight that our mechanistic understanding of the allosteric regulation of the epidermal growth factor receptor is still evolving, and the work presented in Chapters Two and Three present progress towards understanding modes of EGFR allosteric regulation that may have lessons important for understanding how cell surface receptors may be allosterically regulated, more broadly. The second focus reveals mechanisms by which therapeutic antibodies can target oncogenic forms of the EGFR or variants that cause resistance to existing monotherapies. My work in this area has implications for understanding mechanisms of specificity and plasticity of targeting antibodies to cell surface receptor molecules.

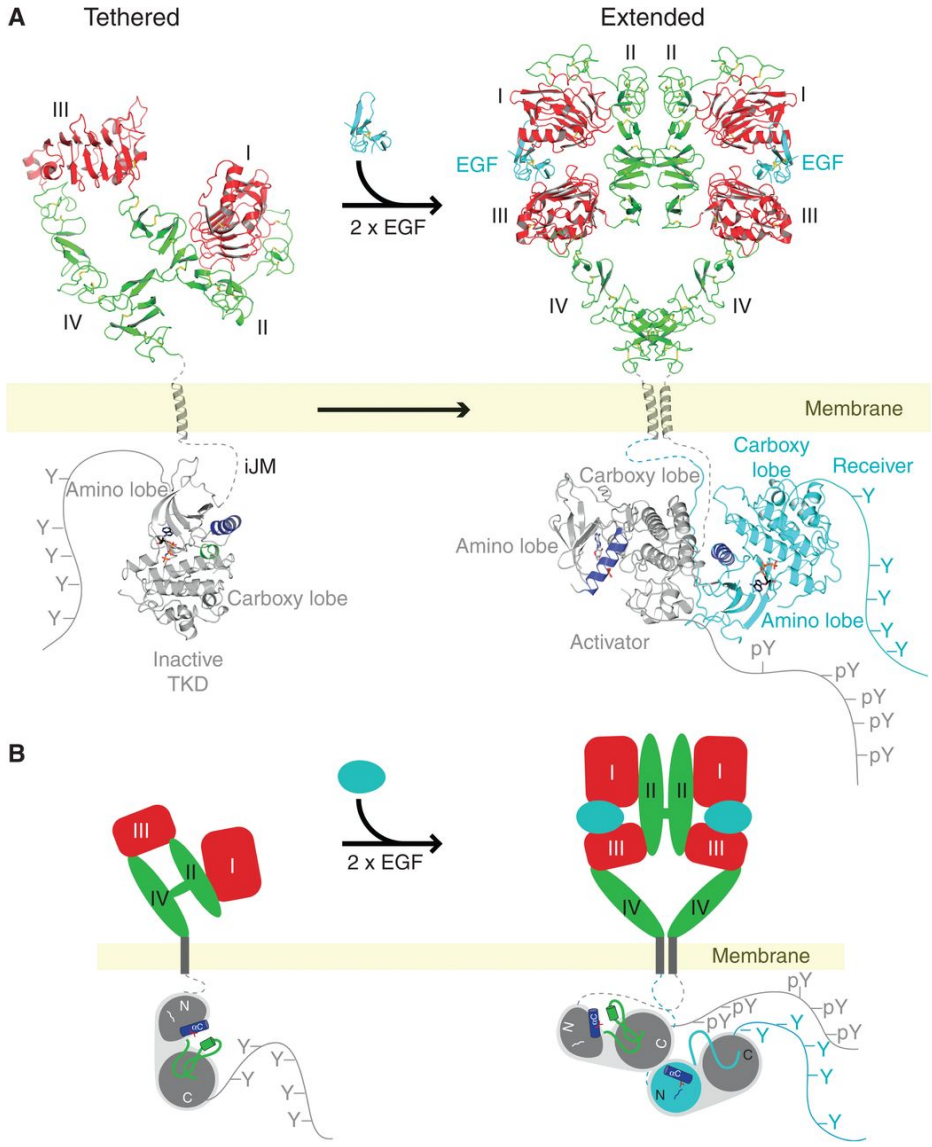


Figure 1.1. Structural Model of Ligand-Regulation of the Epidermal Growth Factor receptor, from (Lemmon, 2014). (A) X-ray structure based model of EGF receptor activation by EGF. In the absence of EGF, the EGFR extracellular region forms an intramolecular 'tether' between two cysteine rich domains II and IV (shown in green). In this

conformation, the ligand binding domains I and III and held too far apart to simultaneously bind ligand. This autoinhibited 'tethered' conformation is presumably linked to an inactive conformation of the tyrosine kinase domain, shown on the bottom part of the left side of panel A. EGF induces a dramatic conformational rearrangement that remodels the conformation of domain II and releases the intramolecular 'tether,' unshielding a β hairpin in domain II that is critical for and drives dimerization of the EGFR extracellular region. Formation of this dimer is indispensable for EGFR activation, as mutation of this hairpin abrogates EGFR activation. This symmetric dimer is linked to an asymmetric dimer of the tyrosine kinase domain, shown on the bottom right side of panel A. Kuriyan and colleagues showed in 2006 that formation of this asymmetric dimer is essential for EGFR activation (Zhang, et al. 2006) . In this asymmetric complex, one kinase monomer in a catalytically inactive conformation allosterically activates the active conformation of the other through N-love/C-lobe interactions. **(B)** A cartoon representation of the process shown in A, highlighting the 'tethered' to 'extended' conformational transition in driving EGFR dimerization.

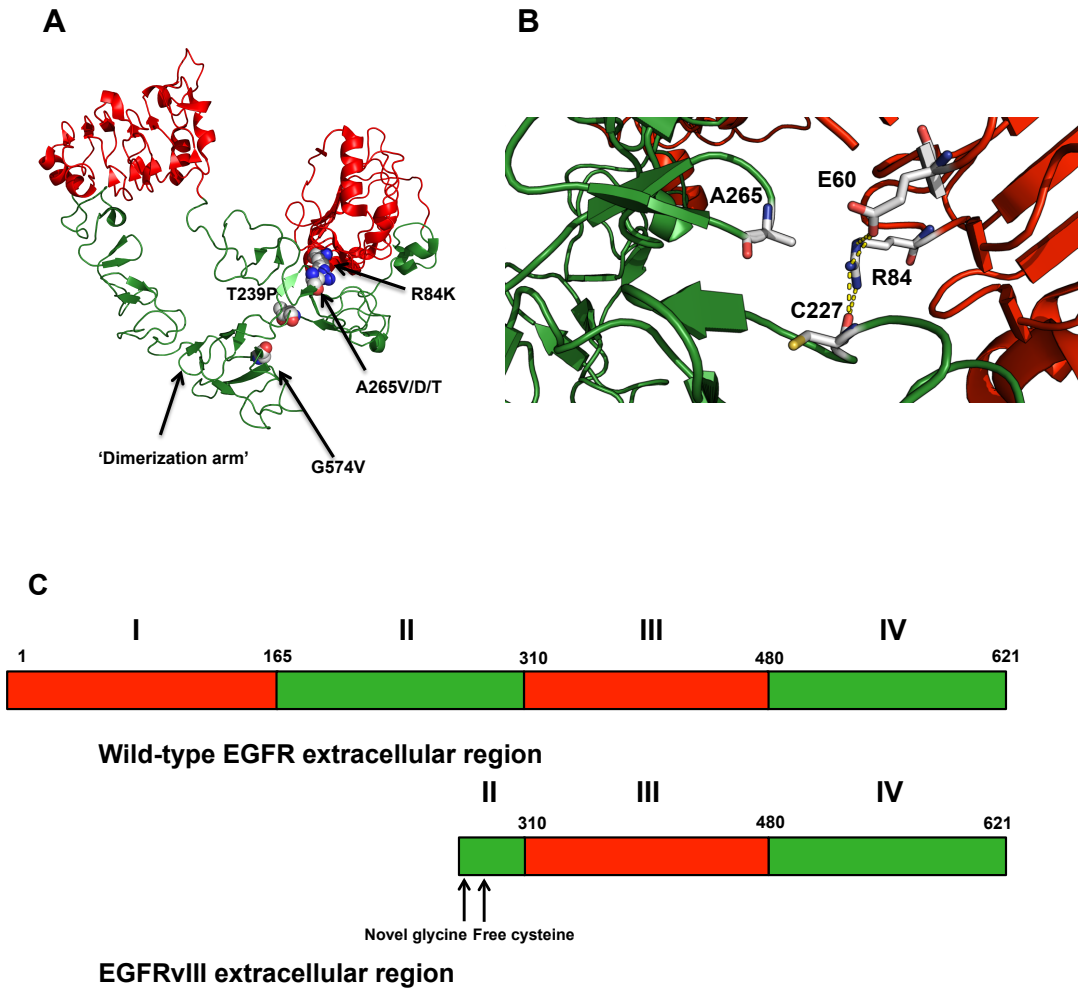


Figure 1.2. Extracellular transforming variants of the Epidermal Growth Factor receptor.

(A) Structural location of identified extracellular point activating and transforming point mutations in EGFR, found in glioblastoma multiforme (GBM) patients or cell lines (Lee et al., 2006). These mutations are largely located in a relatively hydrophobic interface between domains I and II, stabilizing this interface. **(B)** Another common oncogenic variant of EGFR found in ~30-50% of GBM patients is EGFRvIII. EGFRvIII is an gene rearrangement that results in exclusion of exons 2-7 of EGFR, leading to a deletion of mature amino acids 6-273,

and insertion of a unique glycine at its N-terminus, at the exon junction. This mutation also leaves an unpaired cysteine residue in the extracellular region of EGFRvIII, at mature amino acid position 283 (C16 in EGFRvIII).

REFERENCES

Alvarado, D., Klein, D. E., and Lemmon, M. A. (2010). Structural Basis for Negative Cooperativity in Growth Factor Binding to an EGF Receptor. Cell *142*, 568-579.

Alvarado, D., Klein, D.E., and Lemmon, M.A. (2009). ErbB2 resembles an autoinhibited invertebrate epidermal growth factor receptor. Nature *461*, 287-291.

Bertotti, A., Papp, E., Jones, S., Adleff, V., Anagnostou, V., Lupo, B., Sausen, M., Phallen, J., Hruban, C.A., Tokheim, C., Niknafs, N., Nesselbush, M., Lytle, K., Sassi, F., Cottino, F., Migliardi, G., Zanella, E.R., Ribero, D., Russolillo, N., Mellano, A., Muratore, A., Paraluppi, G., Salizzoni, M., Marsoni, S., Kragh, M., Lantto, J., Cassingena, A., Li, Q.K., Karchin, R., Scharpf, R., Sartore-Bianchi, A., Siena, S., Diaz, L.A. Jr., Trusolino, L., Velculescu, V.E. (2015). The genomic landscape of response to EGFR blockade in colorectal cancer. Nature *526*, 263-267.

Bessman, N. J., Bagchi, A., Ferguson, K.M., & Lemmon, M.A. (2014). Complex relationship between ligand binding and dimerization in the epidermal growth factor receptor. Cell Reports *9*, 1306-1317.

Bouyain, S., Longo, P.A., Li, S., Ferguson, K.M., Leahy, D.J. (2005). The extracellular region of ErbB4 adopts a tethered conformation in the absence of ligand. PNAS *102*, 15024-15029.

Burgess, A. W., Cho H.S., Eigenbrot C., Ferguson K.M., Garrett T.P., Leahy D.J., Lemmon M.A., Sliwkowski M.X., Ward C.W., Yokoyama S. (2003). An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. Mol Cell *12*, 541-552.

Carpenter, G., King, L., Jr., Cohen, S. (1978). Epidermal growth factor stimulates phosphorylation in membranes. Nature *276*, 409-410.

Cho, H. S., & Leahy, D.J. (2002). Structure of the extracellular region of HER3 reveals an intermolecular tether. Science *297*, 1330-1333.

Cho, H. S., Mason, K., Ramyar, K.X., Stanley, A.M., Gabelli, S.B., Denney, D.W. Jr. & Leahy, D.J. (2003). Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. Nature *421*, 756-760.

Choi, B. D., Kuan, C.-T., Cai, M., Archer, G.E., Mitchell, D.A., Gedeon, P.C., Sanchez-Perez, L., Pastan, I., Bigner, D.D., & Sampson, J.H. (2013). Systemic administration of a bispecific antibody targeting EGFRvIII successfully treats intracerebral glioma. Proc Natl Acad Sci USA *110*, 270-275.

Chung, I., Akita, R., Vandlen, R., Toomre, D., Schlessinger, J., Mellman, I. (2010). Spatial control of EGF receptor activation by reversible dimerization on living cells. Nature *464*, 783-787.

Cote, G. M., Sawyer, D.B., Chabner, B.A. (2012). ERBB2 inhibition and heart failure. N Engl J Med *367*, 2150-2153.

Dawson, J. P., Berger, M. B., Lin, D., Schlessinger, J., Lemmon, M. A., Ferguson, K. M. (2005). EGF receptor dimerization and activation require ligand-induced conformational changes in the dimer interface. Mol Cell Biol *25*, 7734-7742.

Dawson, J. P., Bu, Z., and Lemmon, M.A. (2007). Ligand-induced structural transitions in ErbB receptor extracellular domains. Structure *15*, 942-954.

Ferguson, K. M., Berger, M., Mendrola, J., Cho, H.S., Leahy, D.J., and Lemmon, M.A. (2003). EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. Molecular Cell *11*, 507-517.

Ferguson, K. M., Darling, P. J., Mohan, M. J., Macatee, T. L., Lemmon, M. A. (2000). Extracellular domains drive homo- but not hetero-dimerization of erbB receptors. EMBO J *19*, 4632-4643.

Franklin, M. C., Carey, K.D., Vajdos, F.F., Leahy, D.J., de Vos, A.M., & Sliwkowski, M.X. (2004). Insights into ErbB signaling from the structure of the ErbB2-pertuzumab complex. Cancer Cell *5*, 317-328.

Freed, D. M., Alvarado, D.A., & Lemmon, M.A. (2015). Ligand regulation of a constitutively dimeric EGF receptor. Nature Communications *6*, 7380.

Gan, H. K., Kaye, A.H., and Luwor, R.B. (2009). The EGFRvIII variant in glioblastoma multiforme. J Clin Neurosci *16*, 748-754.

Gazdar, A. F. (2009). Activating and resistance mutations of EGFR in non-small-cell lung cancer: role in clinical response to EGFR tyrosine kinase inhibitors. Oncogene *28*, S24-S31.

Hubbard, S. R. (2013). The insulin receptor: both a prototypical and atypical receptor tyrosine kinase. Cold Spring Harb Perspect Biol *5*, aa008946.

Jaiswal, B. S., Kljavin, N. M., Stawiski, E. W., Chan, E., Parikh, C., Durinck, S., Chaudhuri, S., Pujara, K., Guillory, J., Edgar, K. A.*, et al.* (2013). Oncogenic ERBB3 mutations in human cancers. Cancer Cell *23*, 603-617.

Johnson, L. A., Scholler, J., Ohkuri, T., Kosaka, A., Patel, P.R., McGettigan, S.E., Nace, A.K., Dentchev, T., Thekkat, P., Loew, A., Boesteanu, A.C., Cogdill, A.P., Chen, T., Fraietta, J.A., Kloss, C.C., Posey, A.D., Jr., Engels, B., Singh, R., Ezell, T., Idamakanti, N., Ramones, M.H., Li, N., Zhou, L., Plesa, G., Seykora, J.T., Okada, H., June, C.H., Brogdon, J.L., & Maus, M.V. (2015). Rational development and characterization of humanized anti-EGFR variant III chimeric antigen receptor T cells for glioblastoma. Sci Transl Med *7*, 275ra222.

Jura, N., Endres, N.F., Engel, K., Deindl, S., Das, R., Lamers, M.H., Wemmer, D.E., Zhang, X., and Kuriyan, J. (2009). Mechanism for activation of the EGF receptor catalytic domain by the juxtamembrane segment. Cell *137*, 1293-1307.

Kumagai, T., Davis, J. G., Horie, T., O'Rourke, D. M., and Greene, M. I. (2001). The role of distinct p185neu extracellular subdomains for dimerization with the epidermal growth factor (EGF) receptor and EGF-mediated signaling. Proc Natl Acad Sci U S A *98*, 5526-5531.

Lawrence, M. C., Colman, P.M. (1993). Shape complementarity at protein/protein interfaces. J Mol Biol *234*, 946-950.

Lee, J. C., Vivanco, I., Beroukhim, R., Huang, J. H., Feng, W. L., DeBiasi, R. M., Yoshimoto, K., King, J. C., Nghiemphu, P., Yuza, Y.*, et al.* (2006). Epidermal growth factor receptor activation in glioblastoma through novel missense mutations in the extracellular domain. PLoS Med *3*, e485.

Lemmon, M. A., Bu, Z., Ladbury, J. E., Zhou, M., Pinchasi, D., Lax, I., Engelman, D. M., Schlessinger, J. (1997). Two EGF molecules contribute additively to stabilization of the EGFR dimer. EMBO J *16*, 281-294.

Lemmon, M. A., and Schlessinger, J. (2010). Cell signaling by receptor tyrosine kinases. Cell *141*, 1117-1134.

Lemmon, M. A., Schlessinger, J., & Ferguson, K.M. (2014). The EGFR family: not so prototypical receptor tyrosine kinases. Cold Spring Harb Perspect Biol *6*, a020768.

Li, S., Kussie, P., Ferguson, K.M. (2008). Structural basis for EGF receptor inhibition by the therapeutic antibody IMC-11F8. Structure *16*, 216-227, doi: 210.1016/j.str.2007.1011.1009.

Li, S., Schmitz, K.R., Jeffrey, P.D., Wiltzius, J.J., Kussie, P., Ferguson, K.M. (2005). Structural basis for inhibition of the epidermal growth factor receptor by cetuximab. Cancer Cell *7*, 301-311.

Liu, P., Bouyain, S., Eigenbrot, C., and Leahy, D.J. (2012). The ErbB4 extracellular region retains a tethered like conformation in the absence of the tether. Protein Science *21*, 152-155.

Lu, C., Mi, L. Z., Schurpf, T., Walz, T., and Springer, T. A. (2012). Mechanisms for kinase-mediated dimerization of the epidermal growth factor receptor. J Biol Chem *287*, 38244-38253.

Lynch, T. J., Bell, D.W., Sordella, R., Gurubhagavatula, S., Okimoto, R.A., Brannigan, B.W., Harris, P.L., Haserlat, S.M., Supko, J.G., Haluska, F.G., et al. (2004). Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med *350*, 2129-2139.

MacDonald, J. L., Pike, L.J. (2008). Heterogeneity in EGF-binding affinities arises from negative cooperativity in an aggregating system. Proc Natl Acad Sci U S A *105*, 112-117.

Maina, F., Casagranda, F., Audero, E., Simeone, A., Comoglio, P.M., Klein, R., & Ponzetto, C. (1996). Uncoupling of Grb2 from the Met receptor in vivo reveals complex roles in muscle development. Cell *87*, 531-542.

Mattoon, D., Klein, P., Lemmon, M.A., Lax, I. Schlessinger, J. (2004). The tethered configuration of the EGF extracellular domain exerts only a limited control of receptor function. Proceedings of the National Academy of Sciences *101*, 923-928.

McDonell, L. M., Kernohan, K.D., Boycott, K.M., & Sawyer, S.L. (2015). Receptor tyrosine kinase mutations in developmental syndromes and cancer: two sides of the same coin. Hum Mol Genet *24*, R60-66.

Montagut, C., et al. (2012). Identification of a mutation in the extracellular domain of the Epidermal Growth Factor Receptor conferring cetuximab resistance in colorectal cancer. Nature Medicine *18*, 221-223, doi:210.1038/nm.2609.

Ogiso, H., Ishitani, R., Nureki, O., Fukai, S., Yamanaka, M., Kim, J. H., Saito, K., Sakamoto, A., Inoue, M., Shirouzu, M., Yokoyama, S. (2002). Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. Cell *110*, 775-787.

Okamoto, I., Kenyon, L.C., Emlet, D.R., Mori, T., Sasaki, J., Hirosako, C., Ichikawa, Y., Kishi, H., Godwin, A.K., Yoshioka, M., Suga, M., Matsumoto, M., and Wong, A.J. (2003). Expression of constitutively activated EGFRvIII in non-small cell lung cancer. Cancer Sci *94*, 50-56.

Paez, J. G., Janne, P.A., Lee, J.C., Tracy, S., Greulich, H., Gabriel, S., Herman, P., Kaye, F.J., Lindeman, N., Boggon, T.J., et al. (2004). EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science *304*, 1497-1500.

Pike, L. J. (2012). Negative co-operativity in the EGF receptor. Biochem Soc Trans *40*, 15-19.

Price, T. J., et al. (2015). Prevalence and outcomes of patients with EGFR S492R ectodomain mutations in ASPECCT: Panitumumab vs. cetuximab in patients with chemorefractory wild-type KRAS exon 2 metastatic colorectal cancer (mCRC). J Clin Oncol *33*, suppl. 3, abstract 740.

Red Brewer, M., Choi, S.H., Alvarado, D., Moravcevic, K., Pozzi, A., Lemmon, M.A., and Carpenter, G. (2009). The juxtamembrane region of the EGF receptor functions as an activation domain. Molecular Cell *34*, 641-651.

Schmeidel, J., Blaukat, A., Li, S., Knochel, T., & Ferguson, K.M. (2008). Matuzumab binding to EGFR prevents the conformational rearrangement required for dimerization. Cancer Cell *13*, 365-373.

Schmitz, K. R., & Ferguson, K.M. (2009). Interaction of antibodies with ErbB receptor extracellular regions. Exp Cell Res *315*, 659-670.

Shan, Y., Eastwood, M.P., Zhang, X., Kim, E.T., Arkhipov, A., Dror, R.O., Jumper, J., Kuriyan, J., & Shaw, D.E. (2012). Oncogenic mutations counteract intrinsic disorder in the EGFR kinase and promote receptor dimerization. Cell *149*, 860-870.

Sok, J. C., Coppelli, F.M., Thomas, S.M., Lango, M.N., Xi, S., Hunt, J.L., Freilino, M.L., GRaner, M.W., Wikstrand, C.J., Bigner, D.D., Gooding, W.E., Furnari, F.B., and Grandis, J.R. (2006). Mutant epidermal growth factor receptor (EGFRvIII) contributes to head and neck cancer growth and resistance to EGFR targeting. Clinical Cancer Research *12*, 5064.

Sugawa, N., Ekstrand, A.J., James, C.D., & Collins, V.P. (1990). Identical splicing of aberrant epidermal growth factor receptor transcripts from amplified rearranged genes in human glioblastomas. Proc Natl Acad Sci USA *87*, 8602-8606.

Walker, F., Orchard, S.G., Jorissen, R.N., Hall., N.E., Zhang, H.H., Hoyne, P.A., Adams, T.E., Johns, T.G., Ward, C., Garrett, T.P., Zhu, H.J., Nerrie, M., Scott, A.M., Nice, E.C., and Burgess, A.W. (2004). CR1/CR2 interactions modulate the functions of the cell surface epidermal grwoth factor receptor. J Biol Chem *279*, 22387-22398.

Warmuth, M., Kim, S., Gu, X.J., Xia, G., & Adrián, F. (2007). Ba/F3 cells and their use in kinase drug discovery. Curr Opin Oncol *19*, 55-60.

Wong, A. J., Ruppert, J.M., Bigner, S.H., Grzeschik, C.H., Humphrey, P.A., Bigner, D.S., & Vogelstein, B. (1992). Structural alterations of the epidermal growth factor receptor gene in human gliomas. Proc Natl Acad Sci USA *89*, 2965-2969.

Yarden, Y., Schlessinger, J. (1987). Epidermal growth factor induces rapid, reversible aggregation of the purified epidermal growth factor receptor. Biochemistry *26*, 1443-1451.

Ymer, S. I., Greenall, S.A., Cvrljevic, A., Cao, D.X., Donoghue, J.F., Epa, V.C., Scott, A.M., Adams, T.E., & Johns, T.G. (2011). Glioma Specific Extracellular Missense Mutations in the First Cysteine Rich Region of Epidermal Growth Factor Receptor (EGFR) Initiate Ligand Independent Activation. Cancers (Basel) *3*, 2032-2049.

Yu, X., Sharma, K.D., Takahashi, T., Iwamoto, R., and Mekada, E. (2002). Ligand independent dimer formation of epidermal growth factor receptor (EGFR) is a step separable from ligand-induced EGFR signaling. Mol Cell Biol *13*, 2547-2557.

Zhang, X., Gureasko, J., Shen, K., Cole, P.A., Kuriyan, J. (2006). An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. Cell *125*, 1137-1149.

CHAPTER TWO

Complex relationship between ligand binding and dimerization in the Epidermal Growth Factor Receptor

SUMMARY

The EGFR plays pivotal roles in development and is mutated or overexpressed in several cancers. Despite recent advances, the complex allosteric regulation of EGFR remains incompletely understood. In efforts to understand why the negative cooperativity observed for intact EGFR is lost when its isolated extracellular region (ECR) is studied, we uncovered unexpected relationships between ligand binding and receptor dimerization. The two processes appear to compete. Surprisingly, dimerization does not enhance ligand binding (although ligand binding promotes dimerization). We further show that simply forcing EGFR ECRs into pre-formed dimers in the absence of ligand yields ill-defined, heterogeneous structures. Finally, we demonstrate that extracellular EGFR-activating mutations in glioblastoma patients enhance ligand binding affinity without directly promoting EGFR dimerization – suggesting that these oncogenic mutations alter the allosteric linkage between dimerization and ligand binding. Our findings have important implications for understanding how EGFR and its relatives are activated by specific ligands and pathological mutations.

Note: This chapter is reproduced as published in

Bessman, N.J., Bagchi, A., Ferguson, K.M., Lemmon, M.A. Complex relationship between ligand binding and dimerization in the Epidermal Growth Factor Receptor. *Cell Reports*. **9**, 1306-17.

INTRODUCTION

X-ray crystal structures determined in 2002 and 2003 (Burgess, 2003) provided mechanistic pictures of how extracellular ligand binding induces EGFR dimerization. A single ligand binds simultaneously to both domains I and III in the extracellular region (ECR) of one EGFR molecule, driving it into an 'extended' conformation as shown in Figure 2.1. This conformational change exposes a dimerization interface in domain II, through which the ECR then self-associates with a dissociation constant (K_D) in the micromolar range (Burgess, 2003; Dawson, 2005; Dawson et al., 2007). The model depicted in Figure 2.1 satisfyingly explains ligand-induced EGFR dimerization. It fails, however, to capture the complex ligand-binding characteristics seen for EGFR at the cell surface – with concave-up Scatchard plots interpreted either as negative cooperativity (De Meyts, 2008; MacDonald, 2008) or distinct affinity-classes of EGF-binding site (Schlessinger, 1986) with high-affinity sites responsible for EGFR signaling (Defize et al., 1989). This cooperativity or heterogeneity is lost when the ECR from EGFR is studied in isolation, as also observed for the insulin receptor (De Meyts and Whittaker, 2002).

Insight into structural origins of EGF/EGFR binding complexity was provided by studies of the *D. melanogaster* EGFR (dEGFR), which – unlike its human counterpart – retains its negative cooperativity when the soluble ECR is isolated (Alvarado et al., 2010). Crystal structures of the ECR from dEGFR revealed a relatively simple 'half-of-the-sites' reactivity in which occupying one binding site in an asymmetric dEGFR dimer restrains and reduces the ligand-binding affinity of the second site (Alvarado et al., 2010). Subsequent detailed comparisons of human and *Drosophila* receptor ECR dimer structures led Leahy and colleagues to perform experiments suggesting similar half-of-the-sites reactivity for human EGFR (Liu et al., 2012). Moreover, detailed studies of EGF binding to intact hEGFR in cells are consistent with this model (MacDonald, 2008). The observed (or inferred) negative cooperativity requires formation of a stable singly-liganded receptor dimer, a species that is

never seen for the isolated ECR from hEGFR (Lemmon, 1997) but forms readily for its *Drosophila* counterpart (Alvarado et al., 2010). The ECR of the *Drosophila* receptor even dimerizes significantly (K_D ~ 40 μM) without bound ligand (Alvarado, 2009), reminiscent of the ligand-independent (pre-formed) dimers reported for intact, cell-surface, hEGFR in many studies (Lemmon, 2009). We therefore reasoned that artificially dimerizing the ECR from the human receptor might restore negative cooperativity and provide avenues for studying details of the complex ligand-binding characteristics of hEGFR. Indeed, engineered dimers of the hEGFR ECR were previously reported to have increased ligand-binding affinity and concave-up Scatchard plots (Adams et al., 2009; Jones, 1999). Similarly, concave-up Scatchard plots (suggesting negative cooperativity) could be restored to the insulin receptor ECR by fusing it to a dimeric immunoglobulin Fc domain (Bass et al., 1996) or to a dimerizing leucine zipper (Hoyne et al., 2000).

Here, we describe studies of the ECR from human EGFR that has been artificially dimerized, or harbors oncogenic mutations found in glioblastoma patients (Lee et al., 2006). Our findings provide new insight into the heterogeneous nature of pre-formed ECR dimers and origins of negative cooperativity. The data also argue that the extracellular structures induced by ligand binding are not 'optimized' for dimerization and, conversely, that dimerization does not optimize the ligand-binding site. Moreover, glioblastoma mutations appear to affect the allosteric linkage between ligand binding and dimerization rather than simply stabilizing EGFR dimers as expected. These studies have important implications for understanding extracellular activating mutations found in ErbB receptors in glioblastoma and other cancers, and also for understanding specificity of ligand-induced ErbB receptor heterodimerization.

RESULTS and DISCUSSION

Pre-dimerizing the EGFR extracellular region has a modest effect on EGF binding affinity

To access pre-formed dimers of the hEGFR extracellular region (sEGFR) experimentally we used two approaches. In one, we fused a dimerizing Fc domain to the very C-terminus of sEGFR (residue 621 of the mature protein) to create sEGFR-Fc. In the second, we fused a dimeric leucine zipper – from the *S. cerevisiae* transcription factor GCN4 – to the sEGFR C-terminus to yield sEGFR-Zip. Both proteins were expressed using a baculovirus expression system and were purified using approaches similar to those used in our earlier studies of sEGFR (see Experimental Procedures). Size exclusion chromatography (SEC) and/or sedimentation equilibrium analytical ultracentrifugation (AUC) confirmed that the two sEGFR fusion proteins are dimeric (Figure 2.7).

To measure K_D values for ligand binding to sEGFR-Fc and sEGFR-Zip, we labeled EGF with Alexa-488 and monitored its binding to sEGFR proteins by measuring fluorescence anisotropy (FA). As shown in Figure 2.2A, EGF binds approximately 10-fold more tightly to the dimeric sEGFR-Fc or sEGFR-Zip proteins than it does to monomeric sEGFR (Table 2.1). The curves obtained for EGF binding to sEGFR-Fc and sEGFR-Zip fit well to simple one-site binding functions, and did not show any sign of negative cooperativity. This suggests that both EGF-binding sites in the dimer are equivalent, or have K_D values that are sufficiently similar to be indiscernible in this assay. Thus, our initial studies did not support the hypothesis that simply dimerizing human sEGFR restores the negatively cooperative ligand binding that is seen for the intact receptor in cells.

One surprise from these data was that forced sEGFR dimerization has only a small (≤10-fold) effect on EGF-binding affinity. Under the conditions of the FA experiments, isolated sEGFR (without zipper or Fc fusion) remains monomeric; the FA assay contains just 60 nM EGF, so the maximum concentration of EGF-bound sEGFR is also limited to 60 nM – which

is more than 20-fold lower than the K_D for dimerization of the EGF/sEGFR complex (Dawson, 2005; Lemmon, 1997). This ≤10-fold difference in affinity for dimeric and monomeric sEGFR seems modest in light of the strict dependence of sEGFR dimerization on ligand binding (Dawson, 2005; Lax et al., 1991; Lemmon, 1997); sEGFR does not dimerize detectably even at millimolar concentrations in the absence of ligand, but ligand binding promotes receptor self-association with a K_D value in the 1 μM range (corresponding to a Gibbs free energy or ΔG for dimerization of -8 kcal/mol). Straightforward linkage of these equilibria should stabilize EGF binding to dimeric sEGFR by up to 8 kcal/mol compared with binding to monomeric sEGFR. This would translate into an approximately ~10^6-fold increase in affinity of ligand for the dimer. The small difference in EGF-binding affinity for dimeric and monomeric sEGFR is also significantly smaller than the 40-100 fold difference typically reported between high-affinity and low-affinity EGF-binding on the cell surface when data are fit to two affinity-classes of binding site (Burgess, 2003; Magun et al., 1980).

Mutations that prevent sEGFR dimerization do not significantly reduce ligand-binding affinity

The small (≤10-fold) increase in ligand-binding affinity caused by sEGFR pre-dimerization led us to question the extent to which domain II-mediated sEGFR dimerization is linked to ligand binding. It is typically assumed that the domain II conformation stabilized upon forming the sEGFR dimer in Figure 2.1C optimizes the positions of domains I and III for EGF binding. To test this hypothesis, we introduced into sEGFR a well-characterized pair of domain II mutations that block dimerization: one at the tip of the dimerization arm (Y251A) and one at its 'docking site' on the adjacent molecule in a dimer (R285S). The resulting double (Y251A/R285S) mutation abolishes both sEGFR dimerization and EGFR signaling (Dawson, 2005; Ogiso, 2002). Importantly, we used isothermal titration calorimetry (ITC) for these studies, where all interacting components are all free in solution. Previous studies

employing surface plasmon resonance (SPR) have indicated that dimerization-defective sEGFR variants bind immobilized EGF with reduced affinity (Dawson, 2005), and we were concerned that this may reflect avidity artifacts – where dimerizing sEGFR binds more avidly than monomeric sEGFR to sensorchip surfaces bearing immobilized EGF.

Surprisingly, our ITC studies showed that the Y251A/R285S mutation has no significant effect on ligand-binding affinity for sEGFR in solution (Table 2.1). These experiments employed sEGFR (with no Fc fusion) at 10 μM – ten times greater than the K_D value for dimerization of ligand-saturated wild-type sEGFR (K_D ~1 μM). Dimerization of wild-type sEGFR ($sEGFR^{wild\text{-}type}$) should therefore be complete under these conditions, whereas the Y251A/R285S-mutated variant ($sEGFR^{Y251A/R285S}$) does not dimerize at all (Dawson, 2005). The K_D value measured for EGF binding to dimeric $sEGFR^{wild\text{-}type}$ was essentially the same (within 2-fold) as that for $sEGFR^{Y251A/R285S}$ (Figure 2.2B,C and Table 2.1), arguing that the favorable ΔG of sEGFR dimerization (-8 kcal/mol) does not contribute significantly (< 0.4 kcal/mol) to enhanced ligand-binding. Similarly, affinities for TGFα binding to $sEGFR^{wild\text{-}type}$ or $sEGFR^{Y251A/R285S}$ are indistinguishable at 80 nM and 82 nM respectively (Table 2.2). These ITC data lead to two important conclusions. First and most importantly, they show that (when avidity effects in SPR are avoided) domain II-mediated dimerization does not significantly enhance ligand binding – implying that there is no positive linkage between ligand binding and sEGFR dimerization. Second, the results force us to revise the interpretation of EGF/sEGFR ITC studies that we published in 1997 (Lemmon, 1997). We previously ascribed the major entropy-driven event (with positive ΔH) to sEGFR dimerization, modeling EGF binding as an enthalpy-driven event based on ITC of EGF binding to isolated domain III (Figure 2.9). The fact that the entropy-driven event is maintained in the absence of sEGFR dimerization refutes this and reveals that EGF binding to the intact ECR is entropy driven, consistent with the associated conformational changes (see Figure 2.9).

Thus, contrary to the ~10^4-10^6 fold enhancement in ligand-binding expected if the sEGFR dimerization and EGF binding equilibria were straightforwardly linked, our data (summarized in Table 1) reveal that blocking sEGFR dimerization has little influence on ligand-binding affinity. Although ΔG for ligand binding is therefore essentially unchanged by the Y251A/R285S mutation, the enthalpy change (ΔH) associated with EGF binding is more favorable (less positive) by 2.0 kcal/mol for $sEGFR^{wild\text{-}type}$ than for $sEGFR^{Y251A/R285S}$. Compensating for this, $T\Delta S$ is less favorable by 1.6 kcal/mol for binding to $sEGFR^{wild\text{-}type}$. Very similar results were obtained for TGFα (Figure 2.9A/B and Table 2.2). Moreover, direct comparison of ITC and FA experiments shows that EGF binds $sEGFR^{wild\text{-}type}$ with very similar affinities under conditions where it does (ITC) and does not (FA) dimerize (Table 2.1 and Figure 2.2A,B) – also arguing for an absence of positive linkage between ligand binding and dimerization. Dimerization of EGF-bound $sEGFR^{wild\text{-}type}$ is essentially complete in our ITC experiments ([EGF/sEGFR] reaches 10 μM) and is essentially absent in FA ([EGF/sEGFR] ≤ 60 nM), yet both studies report similar EGF-binding K_D values. Moreover, even for covalently-dimerized sEGFR-Fc, mutation of the dimerization arm does not impair ligand binding. As shown in Figure 2.2D, surface plasmon resonance (SPR) analysis of EGF binding to wild-type and Y251A/R285S-mutated sEGFR-Fc revealed that this mutation has a negligible effect on EGF-binding affinity in the Fc fusion.

Thermodynamics of EGF binding to sEGFR-Fc

If there is no discernible positive linkage between sEGFR dimerization and EGF binding, why is it that sEGFR-Fc and sEGFR-Zip bind EGF ~10-fold more strongly than wild-type sEGFR in Figure 2.2A? To investigate this, we compared ITC titrations of EGF into sEGFR-Fc and sEGFR-Zip (Figure 2.3A and B) with those for EGF into non-fused, isolated sEGFR. As shown in Table 2.1, the positive ΔH for EGF binding to pre-dimerized sEGFR is greater (more unfavorable) than for binding to isolated (non-dimerized) $sEGFR^{wild\text{-}type}$,

suggesting that enforced dimerization may actually impair ligand/receptor interactions such as hydrogen bonds and salt-bridges. This increased ΔH is more than compensated for, however, by a favorable increase in $T\Delta S$. We suggest that this favorable entropic effect might reflect an 'ordering' imposed upon unliganded sEGFR when it is pre-dimerized, such that it exhibits fewer degrees of freedom compared to monomeric sEGFR. In particular, since EGF binding does induce sEGFR dimerization it seems clear that pre-dimerization will reduce the entropic cost of bringing two sEGFR molecules into a dimer upon ligand binding.

Possible heterogeneity of binding sites in sEGFR-Fc

Closer inspection of titrations of EGF into sEGFR-Fc such as that shown in Figure 2.3A suggested some heterogeneity of sites, as evidenced by the slope in the early part of the experiment. To investigate this possibility further, we repeated the titrations over a range of temperatures. We reasoned that, if there are in fact two different types of EGF-binding site in the sEGFR-Fc dimer, they might be characterized by different values for heat capacity change (ΔC_p). If so, the small differences might become more evident at higher (or lower) temperatures. Indeed, ΔC_p values have been correlated with the nonpolar surface area buried upon binding (Livingstone et al., 1991) – and we know that this differs for the two Spitz-binding sites in the asymmetric *Drosophila* EGFR dimer (Alvarado et al., 2010). As shown in Figure 2.3C, the heterogeneity was more easily discernible at higher temperatures for sEGFR-Fc – especially at 25˚C and 30˚C – suggesting the possible presence of distinct classes of binding sites in the sEGFR-Fc dimer. We were not able to fit the two K_D values (or ΔH values) uniquely with any precision, since the experiment does not have enough information for unique fitting to a model with 4 variables. Whereas binding to wild-type sEGFR could be fit confidently with a single-site binding model throughout the temperature range in Figure 2.3C, enforced sEGFR dimerization (by fusion to Fc) creates some apparent

heterogeneity in binding sites, which may reflect negative cooperativity of the sort seen with dEGFR. The different binding sites are too close in their K_D values to be discerned in ITC or FA studies, and can only be distinguished based on different ΔH values at higher temperature. Nonetheless, these data do suggest that negative cooperativity may indeed be an intrinsic property of the hEGFR extracellular region as suggested (Liu et al., 2012), and as visualized for the *Drosophila* receptor (Alvarado et al., 2010). Presumably, interactions involving other parts of EGFR are responsible for the greater distinction in K_D values seen for the intact receptor (Macdonald-Obermann and Pike, 2009).

Ligand binding is required for well-defined dimerization of the EGFR extracellular region

To investigate the structural nature of the pre-formed ECR dimer in sEGFR-Fc, we used negative stain electron microscopy (EM). We hypothesized that enforced dimerization in the Fc fusion might cause the ligand-free ECR from hEGFR to form the same type of loose domain II-mediated dimer that was seen in crystals of its unliganded *Drosophila* counterpart (Alvarado, 2009). As shown in Figure 2.4A, it was clear in the presence of saturating ligand that the ECRs in sEGFR-Fc form the characteristic heart-shape dimer seen crystallographically and in EM studies by Springer's group (Lu et al., 2010; Mi et al., 2011). Figure 2.4B presents a structural model of the liganded sEGFR dimer fused to a dimeric Fc domain, and Figure 2.4C shows a calculated projection of this model at 12 Å resolution. The selection of class averages for sEGFR-Fc plus EGF in Figure 2.4A resemble this model very closely, confirming that the expected ECR dimer does indeed form. Class averaging of single particles yielded clear densities for all 4 receptor domains, arranged as expected for the EGF-induced domain II-mediated 'back-to-back' extracellular dimer shown in Figure 1 (Garrett, 2002; Lu et al., 2010). In a subset of classes, the Fc-domain also appeared well resolved with good signal, indicating that these particular arrangements of the Fc domain relative to the ECR represent highly populated states – with the Fc domains occupying similar

positions to those of the kinase domain in detergent-solubilized intact receptors (Mi et al., 2011).

By contrast, in the absence of EGF, EM studies of sEGFR-Fc failed to yield signal-enhanced class averages with well-defined and interpretable inter-domain relationships (Figure S4) – despite significant effort with the same protein preparations and staining conditions that yielded the clear averages shown in Figure 2.4A. Thus, simply forcing the ECR from hEGFR into a dimer by fusing to an Fc domain does not cause it to form a well-ordered domain II-mediated back-to-back dimer. Ligand binding is required for this type of dimer to form.

Whether the ECRs are tethered or extended in unliganded sEGFR-Fc (or sample both conformations) is not clear. Solution small angle X-ray scattering (SAXS) studies showed that sEGFR-Fc becomes more significantly compact upon EGF binding, with the radius of gyration (R_g) falling from 65.0 Å to 56.4 Å (Figure 2.4D) and the maximum interatomic distance (D_{max}) within the molecule falling from 198 Å to 175 Å (Figure 2.4E). Values for R_g and D_{max} for ligand-bound sEGFR-Fc agree reasonably well (within 8%) with those calculated for a model with the Fc domain dimer appended to an extended sEGFR dimer resembling that seen in crystal structures of EGF/EGFR complexes (Figure 2.4B and 2.4F, model i) – although flexibility in the relative orientations of the ECR and Fc regions make direct comparison difficult. The relatively large R_g and D_{max} values for sEGFR-Fc in the absence of ligand (Figure 2.4D and 2.4E) are more consistent with a model in which the ECRs are splayed apart, possibly while remaining tethered. Indeed, a model in which two tethered EGFR ECRs are attached to the Fc dimer and maximally splayed apart (Figure 2.4F, model ii) yields an R_g of 69 Å (compared with 65 Å for unliganded sEGFR-Fc) and a D_{max} value of 212 Å (compared with 198 Å for unliganded sEGFR-Fc). An Fc-fused dimer in which the tethered sEGFR moieties are adjacent (Figure 4F, model iii) is more compact, suggesting that unliganded sEGFR-Fc may lie (on average) between models ii and iii. Thus, in agreement with the EM results, our SAXS data also argue that ligand binding is necessary for

formation of the well-defined domain II-mediated dimerization interface. Simply forcing the receptor molecules into close proximity is not sufficient – as Springer and colleagues also concluded in related studies (Lu et al., 2012).

Our results, and those of Lu et al (2012), argue that pre-formed dimers of the human EGFR extracellular region do not contain a well-defined domain II-mediated interface. Rather, the ECRs in these dimers are likely to sample a broad array of positions (and possibly conformations). This conclusion argues against recently published suggestions that stable unliganded extracellular dimers "disfavor activation in pre-formed dimers by assuming conformations inconsistent with" productive dimerization of the rest of the receptor (Arkhipov et al., 2013). The ligand-free inactive dimeric ECR species modeled by Arkhipov et al., 2013 are simply not stable. Indeed, it has been known for some time that the isolated ECR from EGFR has a very low propensity for self-association without ligand, with a K_D value in millimolar range (Lemmon, 1997). The data presented here further show that sEGFR does not form a defined structure even when forced to dimerize in an Fc fusion. It is therefore difficult to envision how the ECR could disfavor activation in pre-formed dimers by assuming any particular conformation. It has also been argued that the unliganded extracellular region somehow impedes dimerization driven by the intracellular region (Endres et al., 2013). The orientational flexibility of the ECR indicated by our EM and SAXS studies of sEGFR-Fc and by studies from the Springer laboratory (Lu et al., 2010; Lu et al., 2012; Mi et al., 2011) makes it very difficult to imagine how the ECR could sterically constrain dimerization mediated by the other parts of EGFR. Moreover, if ligand binding activates EGFR simply by removing steric constraints imposed by the ECR it is difficult to understand why specific mutations of even single residues in the dimerization interface should block activation (Dawson, 2005; Ogiso, 2002), and conversely why mutations that destabilize the tethered configuration are not activating (Mattoon et al., 2004; Walker et al., 2004).

Structural implications of weak linkage between ligand binding and dimerization

It has typically been assumed that binding of a single ligand molecule between domains I and III of an sEGFR molecule stabilizes a structure resembling one half of the 2:2 receptor dimer in Figure 2.1C. Indeed, this is the assumption in Figure 2.1B. In addition, the domain II conformation in a ligand-bound sEGFR monomer is thought to be ideally suited (or poised) for dimerization (Dawson, 2005; Dawson et al., 2007). This assumption predicts (and presumes) that ligand binding and dimerization are strongly positively linked for EGFR and sEGFR. The almost complete lack of such linkage in our studies suggests that the domain II conformation stabilized by EGF binding may not be optimal for dimerization. Indeed, the precise conformation of domain II in a liganded sEGFR monomer is not known, even though SAXS studies of a non-dimerizing sEGFR variant clearly showed that it becomes extended upon EGF binding, with the dimerization arm exposed (Dawson et al., 2007). Our studies of the *Drosophila* EGFR (Alvarado et al., 2010) also showed clearly that restraining the domain II conformation through interactions at the dimerization interface can significantly impair ligand binding affinity (this is the origin of negative cooperativity). With this precedent in mind, we suggest that domain II-mediated sEGFR dimerization may distort the domain II conformation in a way that compromises ligand binding to domains I and III. Conversely, we suggest that the domain II conformation stabilized by ligand binding is suboptimal for dimerization. In this scenario, ligand-binding and dimerization contacts would exert opposite influences on domain II – effectively competing with one another. This competition could effectively nullify the expected positive linkage between ligand binding and dimerization. If this view is correct, the ligand-bound sEGFR dimer visualized by crystallography would reflect a ‘compromise’ in which domain II adopts a structure that is intermediate between the ideal conformation for ligand binding and the ideal conformation for domain II-mediated dimerization.

The precise role played by the extracellular region (ECR) in EGFR activation has been a subject of debate in recent years. The ECR was initially viewed as a module that

serves simply to drive ligand-induced receptor dimerization (Burgess, 2003). Some more recent data also support this view (Lu et al., 2010; Mi et al., 2011), and this is the simple view that predicts positive linkage between ligand binding and dimerization. Alternatively, the ECR has been argued to function as a steric impediment to ligand-independent receptor dimerization, relieved only when the ECR binds ligand (Chantry, 1995; Endres et al., 2013; Jura, 2009) – as mentioned above. In a third possibility it has been proposed that the ligand-bound ECR dimer must achieve a particular conformation in order for the receptor to be active (Alvarado et al., 2010; Arkhipov et al., 2013; Lemmon, 2009; Liu et al., 2012; Wilson et al., 2009). Although ligand binding promotes ECR dimerization – presumably by exposing the dimerization arm as in Figure 2.1 – the compromise between optimal ligand binding and optimal dimerization will result in a particular conformation, which in turn may be required for productive signaling. Different discrete structures (stabilized by different ligands) may even signal differently (Wilson et al., 2009). A similar compromise between dimerization and ligand binding is also seen in our studies with TGFα (Figure 2.9 and Table 2.2), bolstering the view that this conformational 'competition' may be functionally important. There is nothing in our data, however, to support a role for ligand binding in reversing a pre-existing steric hindrance to dimerization.

Extracellular oncogenic mutations observed in glioblastoma alter linkage between ligand binding and sEGFR dimerization

Missense mutations in the extracellular region of hEGFR were discovered in several human glioblastoma multiforme samples or cell lines, and occur in 10-15% of glioblastoma cases (Brennan et al., 2013; Lee et al., 2006). Several were shown to elevate basal receptor phosphorylation and to cause EGFR to transform NIH-3T3 cells in the absence of EGF (Lee et al., 2006). Thus, these are oncogenic mutations that constitutively activate EGFR, although the mutated receptors can be further activated by ligand (Lee et al., 2006; Vivanco et al., 2012). Two of the most commonly mutated sites in glioblastoma, R84 and A265 (R108 and

A289 in pro-EGFR) are in domains I and II of the ECR respectively, and contribute directly to intramolecular interactions between these domains in inactive sEGFR that are thought to be autoinhibitory (Figure 2.5). Domains I and II become separated from one another in this region upon ligand binding to EGFR (Alvarado, 2009), as illustrated in the lower part of Figure 2.5. Interestingly, analogous mutations in the EGFR relative ErbB3 have also been reported in colon and gastric cancers (Jaiswal et al., 2013).

We hypothesized that mutations in the domain I/II interface might activate EGFR by disrupting autoinhibitory interactions between these two domains – possibly promoting a domain II conformation capable of driving dimerization even in the absence of ligand. We used sedimentation equilibrium AUC to test this hypothesis, but found that sEGFR variants harboring R84K, A265D or A265V mutations all remained completely monomeric in the absence of ligand (Figure 2.6A) at a concentration of 10 μM, which is similar to that experienced at the cell surface (Lemmon, 1997). As with wild-type sEGFR, however, adding ligand promoted dimerization of each mutated sEGFR variant – with K_D values that were indistinguishable from that measured for wild-type. Thus, the extracellular EGFR mutations seen in glioblastoma do not simply promote ligand-independent dimerization of the receptor's ECR – consistent with our finding that even dimerized sEGFR-Fc requires ligand binding in order to form the characteristic 'heart-shaped' dimer observed in crystallographic studies.

Interestingly, the ligand-binding affinity of sEGFR was significantly increased by all of the glioblastoma-derived mutations studied here (Figure 2.6B). The effects ranged from a 5-fold increase for A265D to an almost 20-fold increase for R84K. Similar affinity increases were also seen for TGFα binding (Figure 2.11). ITC studies of EGF binding (Figure 2.6C) further revealed that the 5-20 fold increase in ligand-binding affinity (a 1-2 kcal/mole reduction in ΔG) for R84K and A265V variants can be accounted for by ΔH values that are more favorable (less positive) by around 3 kcal/mole when compared to the values for wild-type

sEGFR (Table 2.1). The A265D variant differs (with a less favorable ΔH), possibly because it introduces a charged group into a relatively hydrophobic region of the protein. These data are consistent with a model in which the glioblastoma mutations in the domain I/II interface "free up" domains I and II to occupy positions that permit more optimal interactions with bound ligand. For example, replacement of R84 with a lysine – seemingly a rather conservative substitution – may destabilize the domain I/II interface by disrupting its hydrogen bond network.

We suggest that domain I is normally restrained by domain I/II interactions so that its orientation with respect to the ligand is compromised. When the domain I/II interface is weakened with mutations, this effect is reversed. If this results simply in increased ligand-binding affinity of the monomeric receptor, the biological consequence might be to sensitize cells to lower concentrations of EGF or TGFα (or other agonists). However, cellular studies of EGFR harboring glioblastoma-derived mutations (Lee et al., 2006; Vivanco et al., 2012) suggest ligand-independent activation, arguing that this is not the key mechanism. Instead, the domain I/II interface mutations may also reduce restraints on domain II so as to permit dimerization of a small proportion of intact receptor driven by the documented interactions that promote self-association of the transmembrane, juxtamembrane, and intracellular regions or EGFR (Endres et al., 2013; Lemmon et al., 2014; Red Brewer et al., 2009).

One particularly interesting possibility is that the elevated ligand-binding affinity caused by the R84K, A265V and A265D mutations does not reflect enhanced binding to monomeric sEGFR, but instead reflects some 'rescue' of the expected linkage between ligand binding and dimerization – so that that ligand does bind significantly more strongly to dimers than monomers for these variants. In this model, the domain I/II interface mutations would reduce communication between the dimerization and ligand-binding sites so that optimal domain II-mediated dimerization has less of a restraining influence on the ligand-binding site. The importance of influences on domain II conformation in activation of EGFR by

the glioblastoma mutations is also supported by studies of tether mutations in EGFR that alter ligand binding in almost exactly the same way, but do not activate the receptor (and do not impact domain II). Mutations that disrupt the intramolecular tether seen in Figure 2.1A enhance ligand binding to sEGFR to the same degree as the glioblastoma mutations (Elleman et al., 2001; Ferguson et al., 2003). Moreover, four different types of tether-disrupting mutation all have essentially the same effect on ligand-binding thermodynamics (Figure 2.12) as seen for R84K or A265V (ΔH becomes more favorable by ~4 kcal/mole). None of these tether-disrupting mutations constitutively activates EGFR, however (Mattoon et al., 2004; Walker et al., 2004). The key difference between the (non-activating) tether mutations and the (activating) glioblastoma mutations is that only the latter directly influence domain II conformation, arguing that this is the key to the oncogenicity of R84K and A265V/D mutations.

CONCLUSIONS

Setting out to test the hypothesis that simply dimerizing the EGFR extracellular region is sufficient to recover the negative cooperativity lost when it is removed from the intact receptor, we were led to revisit several central assumptions about this receptor. Our findings suggest three main conclusions. First, we find that enforcing dimerization of the hEGFR extracellular region does not drive formation of a well-defined domain II-mediated dimer that resembles ligand-bound ECRs or the unliganded ECR from *Drosophila* EGFR. Our EM and SAXS data show that ligand binding is necessary for formation of well-defined 'heart-shaped' domain II-mediated dimers. This result argues that the unliganded extracellular dimers modeled by Arkhipov et al. (2012) are not stable, and that it is therefore highly improbable that stable conformations of pre-formed extracellular dimers disfavor receptor activation by assuming conformations inconsistent with productive dimerization of the rest of the receptor (Arkhipov et al., 2013). Recent work from the Springer laboratory employing

kinase inhibitors to drive dimerization of hEGFR (Lu et al., 2012) has also shown that EGF binding is required to form the 'heart-shaped' ECR dimer shown in Figure 2.1C. These findings leave open the question of the nature of the ECR in pre-formed EGFR dimers, but certainly argue that it is unlikely to resemble the crystallographic dimer seen in the case of unliganded *Drosophila* EGFR (Alvarado, 2009) or what has been suggested by computational studies (Arkhipov et al., 2013).

Second, our ITC results suggest that enforcing dimerization of the ECR from hEGFR may restore some of the complexity in ligand binding that is seen for intact hEGFR, but is lost when the soluble ECR is studied in isolation. Some apparent heterogeneity in EGF-binding sites was restored in our ITC studies of the dimeric sEGFR-Fc fusion. However, since the difference in K_D values was too small to be quantitated with confidence, it is clear that simple dimerization cannot fully recapitulate the apparent negative cooperativity seen for the intact receptor. This finding supports arguments that the transmembrane and/or intracellular regions of the receptor also play an important role in defining negative cooperativity in hEGFR (Adak et al., 2011; Macdonald-Obermann and Pike, 2009), and underlines the need to consider cooperation of interactions mediated by all domains within the intact or nearly-intact receptor (Bessman and Lemmon, 2012). It is interesting that the relative contributions of intra- and extra-cellular regions to the allosteric properties of the receptor appear to differ greatly between mammals and insects – where negative cooperativity is recapitulated in the isolated ECR (Alvarado et al., 2010). Similar differences can also be seen between human receptors within a family. For example, whereas the characteristic concave-up Scatchard plots seen for the intact insulin receptor can only be recapitulated for the isolated ECR of that receptor by fusion to a dimerization domain (Bass et al., 1996; Hoyne et al., 2000), the related IGF1 receptor ECR retains negative cooperativity without such modifications (Surinya et al., 2008). Differences in the relative contributions of intra- and extra-cellular regions to precise receptor regulation may have important signaling relevance.

Third, our calorimetric studies of the ECR from hEGFR show that ligand does not bind with substantially higher affinity to dimers than to monomers – despite the fact that ligand binding is required for ECR dimerization. This result argues that ligand binding is required to permit dimerization but that domain II-mediated dimerization compromises, rather than enhances, ligand binding. Assuming flexibility in domain II, we suggest that it serves to link dimerization and ligand binding allosterically. Optimal ligand binding may stabilize one conformation of domain II in the scheme shown in Figure 1 that is then distorted upon dimerization of the ECR – in turn reducing the strength of interactions with the ligand. Such a mechanism would give the appearance of a lack of positive linkage between ligand binding and ECR dimerization, and a good test of this model would be to determine the structure of a liganded sEGFR monomer (which we expect to differ from a half-dimer). This model also suggests a mechanism for selective hetero- over homo- dimerization of certain ErbB receptors. If a ligand-bound EGFR monomer has a domain II conformation that forms heterodimers with the ErbB2 ECR in preference to EGFR homodimers, this could explain several important observations. It could explain reports that ErbB2 is a preferred heterodimerization partner (Graus-Porta et al., 1997), and might also explain why EGF binds more tightly to EGFR in cells where it can form heterodimers with ErbB2 than in cells that lack ErbB2, where only EGFR homodimers can form (Li et al., 2012). Moreover, if different EGFR agonists stabilize slightly different domain II conformations, this view of a flexible domain II as an allosteric link between ligand binding and dimerization suggests hypotheses for how individual ligands might induce subtly different receptor states or select for specific heterodimer signaling complexes (Sweeney and Carraway, 2000; Wilson et al., 2009). Interestingly – as our data with glioblastoma mutations suggest – alterations in the allosteric communication between domain II and the adjacent ligand-binding domain I can also be oncogenic.

EXPERIMENTAL PROCEDURES

sEGFR constructs

Construction of sEGFR variants with mutations in the domain II/IV tether or the dimerization arm was described previously (Dawson, 2005). Glioblastoma-derived mutations were introduced by standard PCR methods. To generate the sEGFR-Fc fusion, the coding sequence for sEGFR$^{wild\text{-}type}$ in pFb was first mutated to introduce an *Fse I* restriction site close to its 3' end. The cDNA encoding the Fc domain from human IgG1 (IMAGE clone 4575935) was purchased from Open Biosystems, and was used as a PCR template to create a fragment containing the Fc domain coding region flanked by *Fse I* and *Not I* sites at the 5' and 3' ends, respectively. This fragment was then cloned into the pFb-sEGFR$^{wild\text{-}type}$ plasmid containing an *Fse I* site, yielding a plasmid encoding residues 1-645 of the pro-EGFR protein, followed by an alanine-glycine linker (introduced by the *Fse I* site), and then by the 231 residue Fc domain and a C-terminal hexa-histidine tag. The sEGFR-Zip pFb plasmid was constructed by first amplifying a fragment encoding the 33-residue coiled-coil domain from yeast GCN4 by a series of four PCR reactions, each extending the length of the fragment at the 3' end. This fragment also contained 20 nucleotides at the 5' end that were complementary to the C-terminus of sEGFR, as well as a hexa-histidine tag and *Not I* site at the 3' end. This fragment was used as a primer to PCR-amplify (from a pFb-sEGFR$^{wild\text{-}type}$ template) a fragment encoding sEGFR followed by the 33-residue coiled-coil domain and a C-terminal hexahistidine tag, which was then ligated into the pFb plasmid. All constructs were verified by DNA sequencing.

Reagents and proteins

EGF and TGFα were purchased from Millipore Inc. Wild-type and mutated sEGFR variants were expressed in Sf9 cells employing a baculovirus system using the pFastbac

(pFb) plasmid (Invitrogen) as described previously (Ferguson, 2000). Details of constructs and of purification procedures are described below.

Binding and dimerization analyses

Binding of ligand to sEGFR proteins and/or sEGFR dimerization was performed using isothermal titration calorimetry (ITC); fluorescence anisotropy (FA); surface plasmon resonance (SPR) or sedimentation equilibrium analytical ultracentrifugation (AUC), as summarized below.

Protein purification

All sEGFR proteins were purified from the conditioned medium of baculovirus-infected Sf9 cells following the same procedure. Briefly, 4 days post-infection, clarified conditioned media was concentrated to ~1 l and diafiltered against 4 l of 25 mM Tris-HCl, pH 8.0, containing 150 mM NaCl (binding buffer). Diafiltered media was applied to a Ni-NTA agarose column (Qiagen), and sEGFR protein was eluted with increasing concentrations of imidazole in binding buffer. Fractions containing sEGFR were dialyzed, buffer-exchanged, or diluted into 25 mM MES, pH 6.0, containing 30 mM NaCl and sEGFR was purified by cation-exchange chromatography using a Source S column (GE Healthcare) with a gradient of 5% - 50% of 25 mM MES, pH 6.0, containing 500 mM NaCl (over 15 column volumes). The sEGFR-containing fractions were concentrated and protein was further purified by size-exclusion chromatography on a Superose 6 column (GE Healthcare) pre-equilibrated in 25 mM HEPES, pH 8.0, containing 150 mM NaCl. Typical yields were 0.2-1 mg/l of Sf9 cells (depending on the sEGFR variant) and protein was >90% pure by Coomassie stained SDS-PAGE.

Isothermal titration calorimetry (ITC)

All ITC experiments were performed at 25˚C (unless specifically stated otherwise) using a MicroCal ITC200 instrument. Ligand and sEGFR proteins were dialyzed overnight into 20 mM HEPES, pH 8.0, containing 150 mM NaCl, and 3.4 mM EDTA. The sEGFR concentration in the calorimeter cell ranged from 8 to 25 μM, and the concentration of ligand in the syringe ranged from 60 to 280 μM. All protein concentrations were determined by measuring absorbance of purified protein at 280 nm and using the extinction coefficient predicted from primary amino acid sequence. A total of 39 μl of ligand (in 2-3 μl aliquots) was injected over the course of each titration. Data from the first (small) injection were discarded to eliminate syringe leakage artifacts. Ligand titrations into receptor-free ITC buffer were performed to determine the heat of ligand dilution, and these heats were subtracted from the ligand-into-receptor titration data. Data were fit to a single-site binding model in the Origin software package. All titrations were performed independently at least three times, and representative titrations are shown, with values for ΔH and other parameters quoted as mean ± standard deviation.

Fluorescence anisotropy (FA)-based binding assays

EGF was labeled with Alexa-488 utilizing a tetrafluorophenyl ester to label primary amines, according to the protocol provided with the Alexa Fluor 488 Protein Labeling Kit from Molecular Probes (Eugene, OR). Labeled EGF (EGF^{488}) was purified away from free label by size-exclusion chromatography on a Superdex Peptide (GE Healthcare) column equilibrated in 20 mM HEPES, pH 8.0, containing 150 mM NaCl. Labeling efficiency was calculated by determining the concentration of purified, labeled EGF from its absorbance at 280 nm, using an extinction coefficient of 18,825 $cm^{-1}M^{-1}$ (as predicted from the primary sequence of EGF) and at 490 nm, using the extinction coefficient of the Alexa-488 label (71,000 $cm^{-1}M^{-1}$). EGF^{488} at 10 nM (for sEGFR-Fc and sEGFR-Zip) or 60 nM (for $sEGFR^{wild\text{-}type}$) was incubated

with varying amounts of sEGFR protein for 30 minutes at room temperature in 20 mM HEPES, pH 8.0, containing 150 mM NaCl. Fluorescence polarization (FP) measurements for each sample were taken on a Beacon instrument at 20˚C. FP values were converted to anisotropy, and binding curves were derived by assuming that the maximal anisotropy response corresponded to $[EGF_{free}]=0$, and that the anisotropy in the absence of receptor corresponded to $[EGF_{free}]=[EGF_{total}]$. The resulting curves were fit to binding models using the GraphPad Prism software. sEGFR-Fc and $sEGFR^{wild\text{-}type}$ binding data were fit to simple single-site binding models, whereas sEGFR-Zip binding data were fit to a model with a Hill coefficient of ~1.9, presumably reflecting the presence of sEGFR-Zip monomers at very low receptor concentrations. Three independent titrations were performed for each receptor variant.

Surface plasmon resonance (SPR)

SPR binding experiments were performed using a Biacore 3000 instrument at 25˚C. EGF and TGFα were immobilized as described (Dawson, 2005), immobilizing 100-150 response units (RUs). sEGFR variants at various concentrations in 25 mM HEPES, pH 8.0, containing 150 mM NaCl, 3 mM EDTA, and 0.005% Nonidet-p20 were injected at a flow rate of 10 µl/min for 10 minutes, which was sufficient to reach equilibrium even at the lowest concentrations. K_D values for binding of sEGFR variants to these surfaces were determined by fitting the equilibrium responses over a range of concentrations to a single-site Langmuir binding equation using GraphPad PRISM 6.0. Surfaces were regenerated using 1 min injections of 10 mM sodium acetate, pH 5.0, containing 1 M sodium chloride. Multiple rounds of regeneration did not impair sEGFR binding. All experiments were repeated independently at least three times.

Analytical ultracentrifugation (AUC)

Samples of sEGFR wild-type and the glioblastoma-derived variants at concentrations of 10 μM, 5 μM, and 2 μM were loaded into 6-hole sample cells in 25 mM HEPES, pH 8.0, containing 150 mM NaCl for sedimentation equilibrium AUC using a Beckman XL-A instrument and an An-Ti 60 analytical rotor at speeds of 6,000, 9,000 and 12,000 r.p.m. at room temperature. For experiments including TGFα, the ligand was present in a 1.2-fold molar excess over sEGFR protein concentration. Data were analyzed using a single species fit in Sedfit (Schuck et al., 2002) and Sedphat (Schuck, 2003), and were also fit to a monomer-dimer association model as described (Dawson, 2005), using the program HeteroAnalysis (UConn Biotechnology Bioservices Center).

Electron microscopy (EM)

For the EGF/sEGFR-Fc complex, 755 individual particles, manually picked from 150 images using the EMAN2 software package (Tang et al., 2007), were grouped into 10 classes by a reference-free alignment procedure in the program Spider (Shaikh et al., 2008). For the sEGFR-Fc protein alone, 2,566 particles from 673 images were grouped into 20 classes by the same reference-free alignment procedure. For Figure 4C, a model for the sEGFR-Fc fusion protein bound to EGF (shown in Figure 4B) was created by manually superposing PDB entries 3NJP (Lu et al., 2010) and the Fc portion of PDB entry 1HZH (Saphire et al., 2001). A 12 Å resolution map was calculated from this model using the molmap command in the UCSF Chimera software package (Pettersen et al., 2004). This low-resolution map was then used to create 2D projections from multiple angles using the makeboxref command in the EMAN2 software suite. The projection angle that best matched our experimental class averaging results is displayed.

Small-angle X-ray scattering (SAXS)

Protein samples were prepared for SAXS at concentrations of 10-20 μM in 25 mM Tris-HCl, pH 8.0, containing 150 mM NaCl, and 40 minute exposures at 4˚C were performed on a Rigaku S-MAX3000 pinhole camera system, with a Rigaku 007HF rotating anode source and a Rigaku 300 mm wire grid ASM DTR 200 detector.

Scattering data were radially-averaged and reduced to two-dimensional plots using the SAXSgui software, and intensity data from buffer exposures were then subtracted out. Radius of gyration (R_g) values were determined from Guinier plots using the Primus software package (Konarev et al., 2003). The maximum interatomic distance (D_{max}) was obtained by examining P(r) curves generated by the Gnom software package (Svergun, 1992). Briefly, for each scattering dataset, P(r) curves were calculated for a range of D_{max} from 100 to 250 Å, in 5 Å increments. D_{max} was determined by identifying the value that gave the best fit to the experimental scattering data.

Electron microscopy (EM)

Receptor samples at a concentration of 2 μg/ml in 25 mM Tris-HCl, pH 8.0, containing 150 mM NaCl were applied to glow-discharged carbon grids and stained with 0.75% uranyl formate. Images were collected on a Tecnai T12 microscope at 67,000x magnification and operating at 120 keV, and were analyzed as described in Methods.

CHAPTER TWO FIGURES

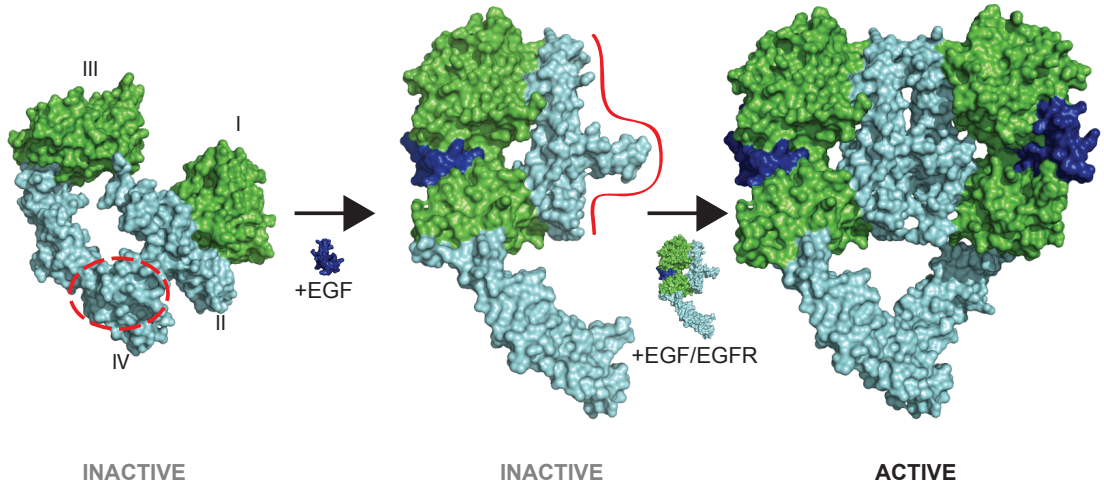


Figure 2.1. Structural view of ligand-induced dimerization of the hEGFR extracellular region (ECR).

(**A**) Surface-representation of tethered, unliganded, sEGFR is shown from PDB entry 1NQL (Ferguson et al., 2003). Ligand-binding domains I and III are colored green, and cysteine-rich domains II and IV are colored cyan. The intramolecular domain II/IV tether is circled in red and labeled. (**B**) Surface representation of a hypothetical model for an extended EGF-bound sEGFR monomer based on SAXS studies showing an extended conformation for an EGF-bound sEGFR variant with a dimerization arm mutation (Dawson et al., 2007). The model was extracted from the extended dimer in PDB entry 3NJP (Lu et al., 2012), but it should be stressed that no details are known for the precise domain II conformation in the extended monomer. The model is colored as in **A**. In addition, EGF is colored blue, the exposed domain II dimerization arm is labeled, and the red boundary represents the primary dimerization interface. (**C**) Surface representation of a 2:2 (EGF/sEGFR) dimer, taken from PDB entry 3NJP (Lu et al., 2012), colored as in **B**. Primary dimerization contacts, mediated by the domain II dimerization arm, are circled in red and labeled. Figure generated by Nicholas Bessman and Mark Lemmon.

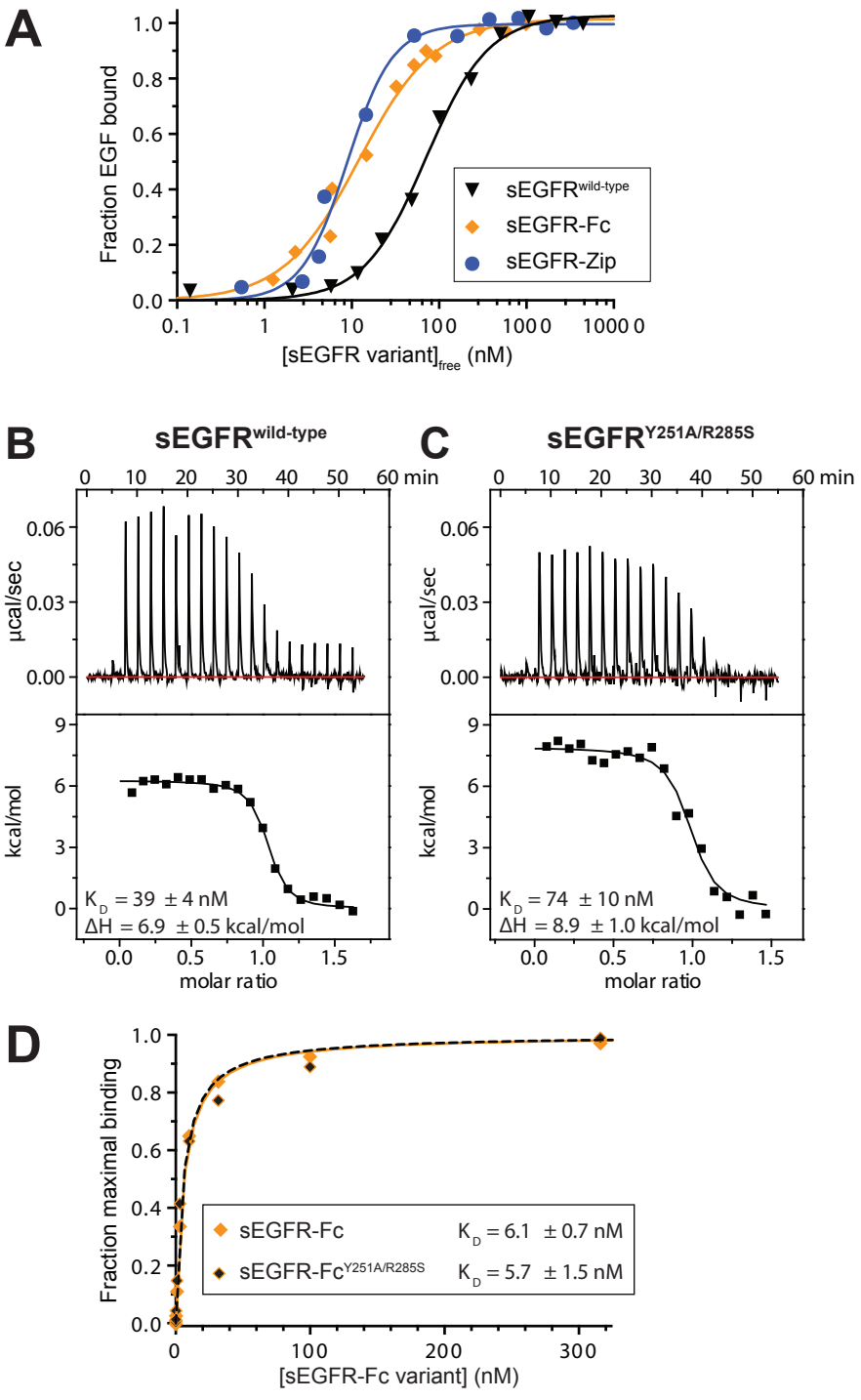


Figure 2.2. Dimerization of the ECR has little effect on its affinity for EGF.

(**A**) Data from fluorescence anisotropy (FA) experiments for binding of Alexa-488-labeled EGF (EGF^{488}) to monomeric $sEGFR^{wild\text{-}type}$ (black triangles), dimeric sEGFR-Fc (orange diamonds) and dimeric sEGFR-Zip (blue circles). Ligand concentrations were 60 nM (for $sEGFR^{wild\text{-}type}$ experiments) or 10 nM (for sEGFR-Fc and sEGFR-Zip). sEGFR-Fc and sEGFR-Zip are both dimeric (see Figure S1), whereas $sEGFR^{wild\text{-}type}$ remains monomeric in this assay. Data shown are representative of three independent experiments. Mean values (± standard deviation) for K_D and other thermodynamic parameters are listed in Table 1. (**B**) Representative ITC analysis of EGF binding to $sEGFR^{wild\text{-}type}$ at 25˚C. EGF (80 μM) was injected into $sEGFR^{wild\text{-}type}$ protein (10 μM) in the ITC cell. Mean K_D and ΔH values (± standard deviation) from three independent experiments are listed (see Table 1). (**C**) ITC analysis of EGF binding to the non-dimerizing $sEGFR^{Y251A/R285S}$ variant. Details are the same for **B**. Equivalent experiments for TGFα binding in Figure S3A and B. (**D**) SPR analysis of EGF binding to constitutively-dimeric sEGFR-Fc, with (solid orange diamond) or without (orange/black diamonds) dimerization-disrupting mutations (Y251A/R285S) in domain II. sEGFR-Fc protein (with wild-type or Y251A/R285S-mutated dimerization interface), was flowed over a sensorchip surface bearing EGF. The equilibrium SPR response for each indicated sEGFR-Fc concentration was measured, and data were fit to a simple single-site binding model. Data are representative of three independent repeats, with best fit K_D values (± standard deviation) noted. Equivalent studies for TGFα are shown in Figure S3C. This figure is the work of Nicholas Bessman. Atrish Bagchi provided some purified proteins used to perform the experiments.

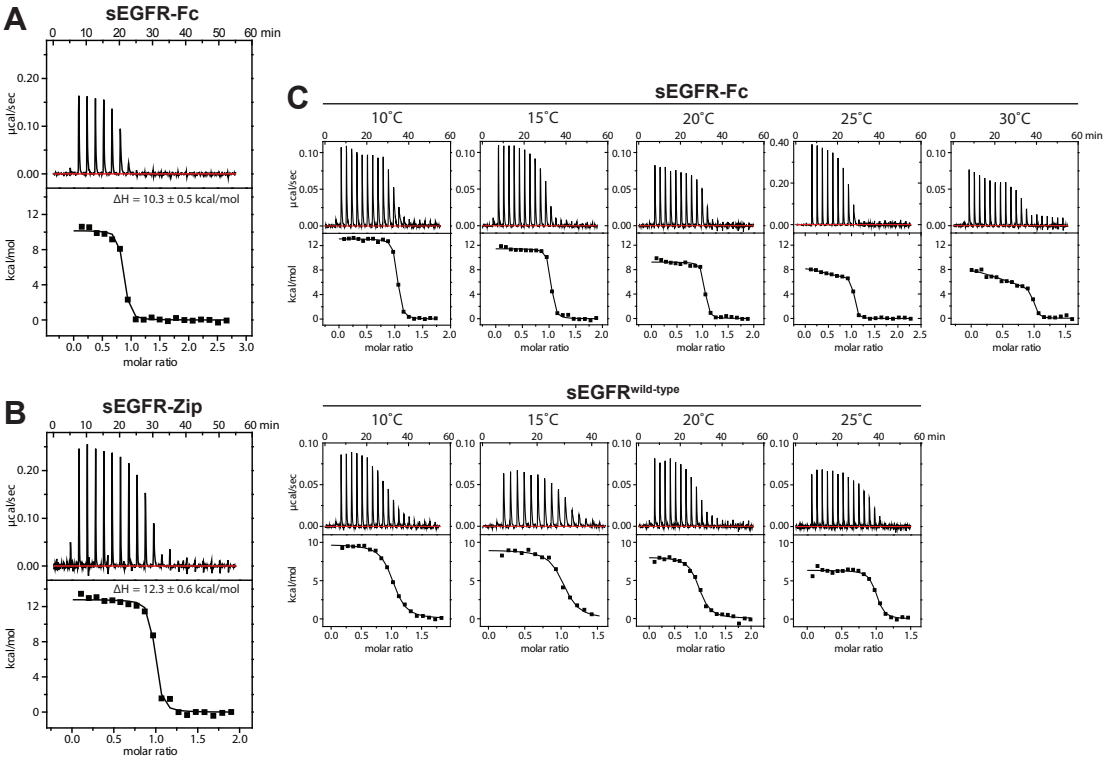


Figure 2.3. Evidence for heterogeneity of sites in forced sEGFR dimers.

(**A**) Representative ITC data for EGF binding to sEGFR-Fc at 25˚C, with EGF at 130 μM in the syringe and sEGFR-Fc in the cell at 8.4 μM. The mean *ΔH* value from three independent experiments (± standard deviation) is listed. (**B**) ITC data for EGF binding to sEGFR-Zip at 25˚C, with EGF in the syringe at 105 μM and sEGFR-Fc in the cell at 11.3 μM. *ΔH* is listed as in **A**, reflecting three independent repeats. (**C**) ITC was performed for EGF binding to sEGFR-Fc (upper graphs) and sEGFR$^{wild\text{-}type}$ (lower graphs) at the different temperatures marked, to resolve potential heterogeneity in ligand binding sites within the dimer. The concentration of EGF in the calorimeter syringe was 80 μM, and the sEGFR protein was present in the cell at 9 μM. For sEGFR-Fc, which showed evidence for more than one binding event, data are shown for an additional titration at 25˚C that used 25 μM sEGFR-Fc in the cell

and 280 μM EGF in the syringe – the improved signal-to-noise confirming existence of a second binding event, potentially reflecting asymmetry in the sEGFR-Fc dimer. This experiment was conceived by Atrish Bagchi and performed by Nicholas Bessman, with protein reagents from Atrish Bagchi. Both authors analyzed data.

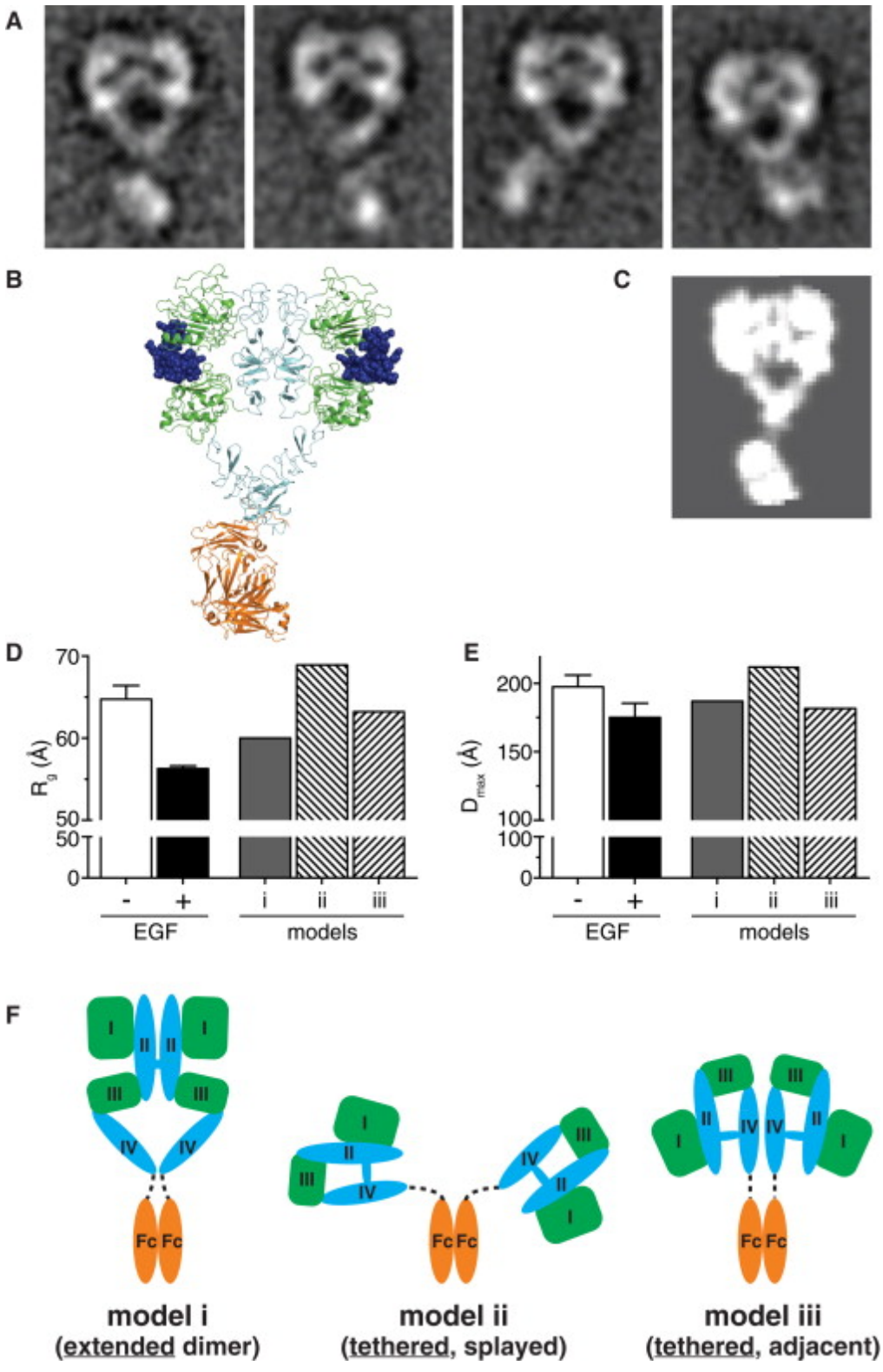

A
B
C
D
70
60
50
50
0
R_g (Å)
- + i ii iii
EGF
models
E
200
150
100
100
0
D_{max} (Å)
- + i ii iii
EGF
models
F
I II II I
III III
IV IV
Fc Fc
model i
(extended dimer)
I II III IV
III IV II I
Fc Fc
model ii
(tethered, splayed)
III III
II IV IV II
I I
Fc Fc
model iii
(tethered, adjacent)

Figure 2.4. Ligand-binding is required for formation of the domain II-mediated 'back-to-back' dimer.

(**A**) Selection of reference-free class averages from single-particle EM images of negatively stained EGF/sEGFR-Fc complexes. Figure S4 shows an example of a field with- and without EGF. (**B**) Structural model for an EGF/sEGFR-Fc complex derived by appending the Fc portion from PDB entry 1HZH (Saphire et al., 2001) to the EGF-bound sEGFR dimer from PDB entry 3NJP (Lu et al., 2012). EGF is colored blue (space filling), the ligand-binding domains I and III in sEGFR are green, cysteine-rich domains II and IV are cyan, and the Fc portion is orange. (**C**) Two-dimensional projection from a calculated 12 Å resolution map based on the EGF/sEGFR-Fc model shown in **B**, generated as described in Supplemental Experimental Procedures. This projection strongly resembles the experimentally-determined class averages for EGF-bound sEGFR-Fc. (**D**) R_g values derived from Guinier analysis of SAXS data for samples containing 10-20 μM sEGFR-Fc in the absence (open bar) or presence (black bar) of a 1.3-fold molar excess of EGF. Experimental data represent the mean of four independent experiments (± standard deviation). R_g values calculated from the three models (i – iii) shown in **F** are also plotted. (**E**) SAXS-derived values for maximum interatomic distance (D_{max}) for sEGFR-Fc alone (open bar) and the EGF/sEGFR-Fc complex (black bar). Data represent means (± standard deviation) from 4 independent repeats. Calculated D_{max} values for the three models shown in **F** are also plotted for comparison. (**F**) Three distinct structural models (i, ii, and iii) were constructed for unliganded sEGFR-Fc. In model i, the sEGFR moiety forms the back-to-back dimer seen in the presence of ligand (or for the unliganded ECR from *Drosophila* EGFR). In models ii and iii, sEGFR retains its tethered conformation, but the two sEGFR moieties in the dimer are either maximally splayed apart (model ii) or adjacent to one another (model iii). R_g and D_{max} values were calculated for each model, for comparison to experimentally-determined parameters. This work was conceived by Mark Lemmon and Nicholas Bessman, and performed by Nicholas Bessman.

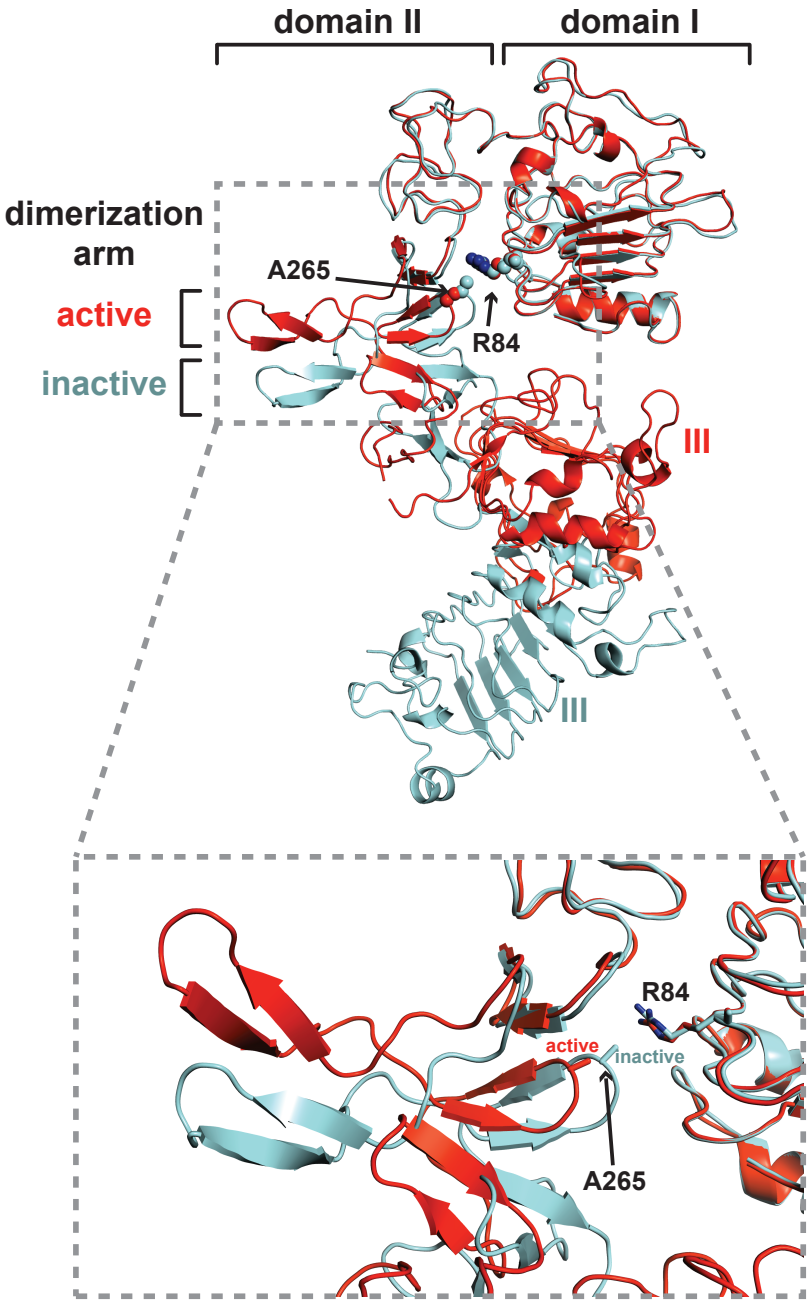


Figure 2.5. Location of EGFR domain I/domain II interface mutations in glioblastoma.

Cartoon representations of sEGFR crystal structures in liganded (red) and unliganded (cyan) states are shown, from PDB entries 1MOX (Garrett, 2002) and 1YY9 (Li et al., 2005) respectively. The two structure are aligned using domain I (residues 1-184) as reference. The side-chains of R84 and A265 are shown, where the majority of mutations have been seen in glioblastoma (Lee et al., 2006; Vivanco et al., 2012). These side-chains are buried in a relatively hydrophobic interface between domains I and II in the unliganded structure. This region of the structure is expanded in the lower part of the figure to show how ligand binding forces domain II away from domain I, disrupting the domain I/II interface. Mutations at R84 and A265 may activate the receptor by mimicking this effect.

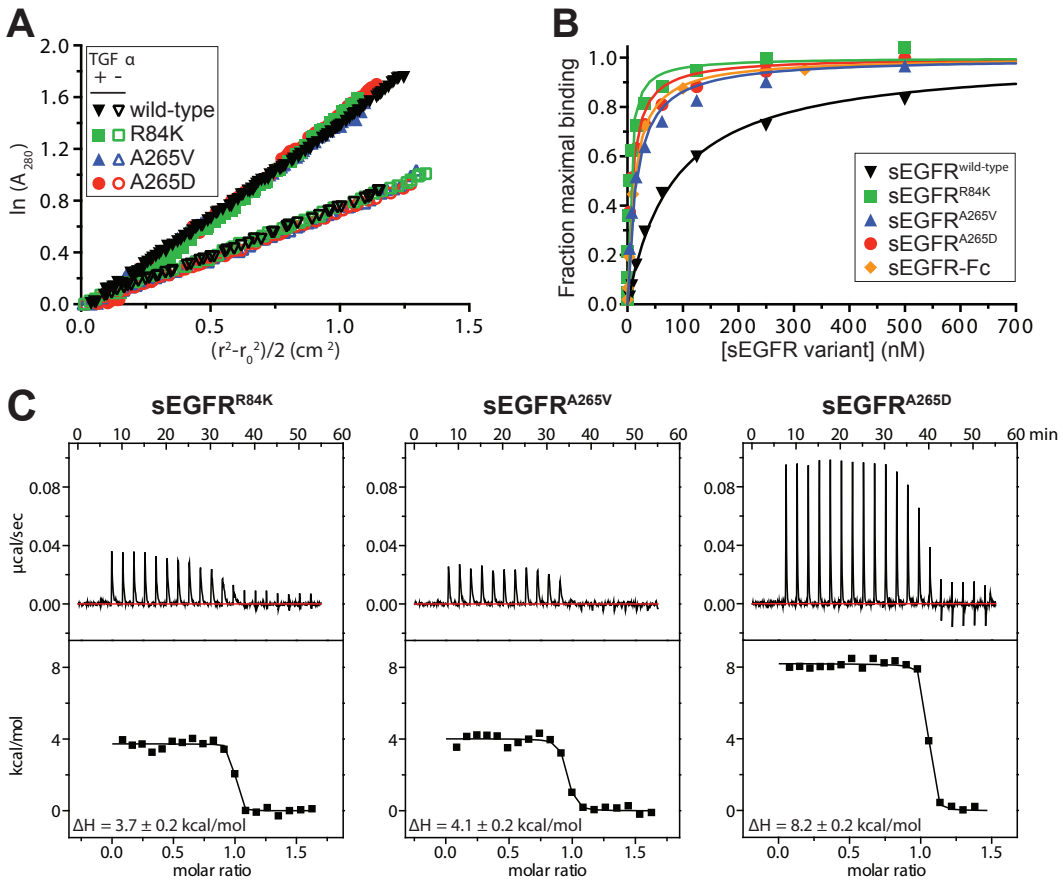


Figure 2.6. Effects of glioblastoma mutations on sEGFR properties.

(**A**) Sedimentation equilibrium AUC analysis of sEGFR variants harboring mutations found in glioblastoma, performed with each sEGFR variant either alone at 10 μM (open symbols) or at 5 μM in the presence of a 1.2-fold molar excess of TGFα (closed symbols). Data are plotted as the natural logarithm of absorbance at 280 nm (A_{280}, monitoring protein concentration) against a function of the radius squared [$(r^2-r_0^2)/2$] for AUC data obtained at 9000 r.p.m. at room temperature, where r is the radial position in the sample and r_0 is the radial position of

the meniscus. For an ideal single species, data transformed in this way yield a straight line with slope proportional to molecular mass (Cantor and Schimmel, 1980). TGFα was used instead of EGF since it contains only one tyrosine (and no tryptophans), so contributes negligibly to A_{280}. In the absence of ligand, best-fit molecular masses were 75 kDa (wild-type); 80 kDa (R84K and A265V); and 89 kDa (A265D). The slight elevation for A265D may reflect some aggregation, but these data argue that the mutations do not cause constitutive dimerization. In the presence of TGFα, single-species fits yielded molecular masses of 157 kDa (wild-type); 141 kDa (R84K); 143 kDa (A265V); and 175 kDa (A265D). Estimated K_D values for sEGFR dimerization in the presence of TGFα, fit as described (Dawson, 2005), were 1.5 μM (R84K); 1.0 μM (A265V); and 4.8 μM (A265D) – compared with 1.2 μM for wild-type sEGFR (**B**) Ligand binding by each mutated sEGFR variant was analyzed using SPR, flowing the protein at a range of concentrations over EGF immobilized on a sensorchip. Best fit K_D values for EGF binding were 83 ± 2.6 nM (wild-type); 4.3 ± 0.9 nM (R84K); 16 ± 2.6 nM (A265V); and 8.6 ± 3.0 nM (A265D). Similar data for TGFα are shown in Figure S5. (**C**) ITC analysis of EGF binding to sEGFR variant harboring mutations found in glioblastoma patients. Titrations were performed with 80 μM EGF aliquots injected into 10 μM sEGFR protein. This figure is the work of Atrish Bagchi.

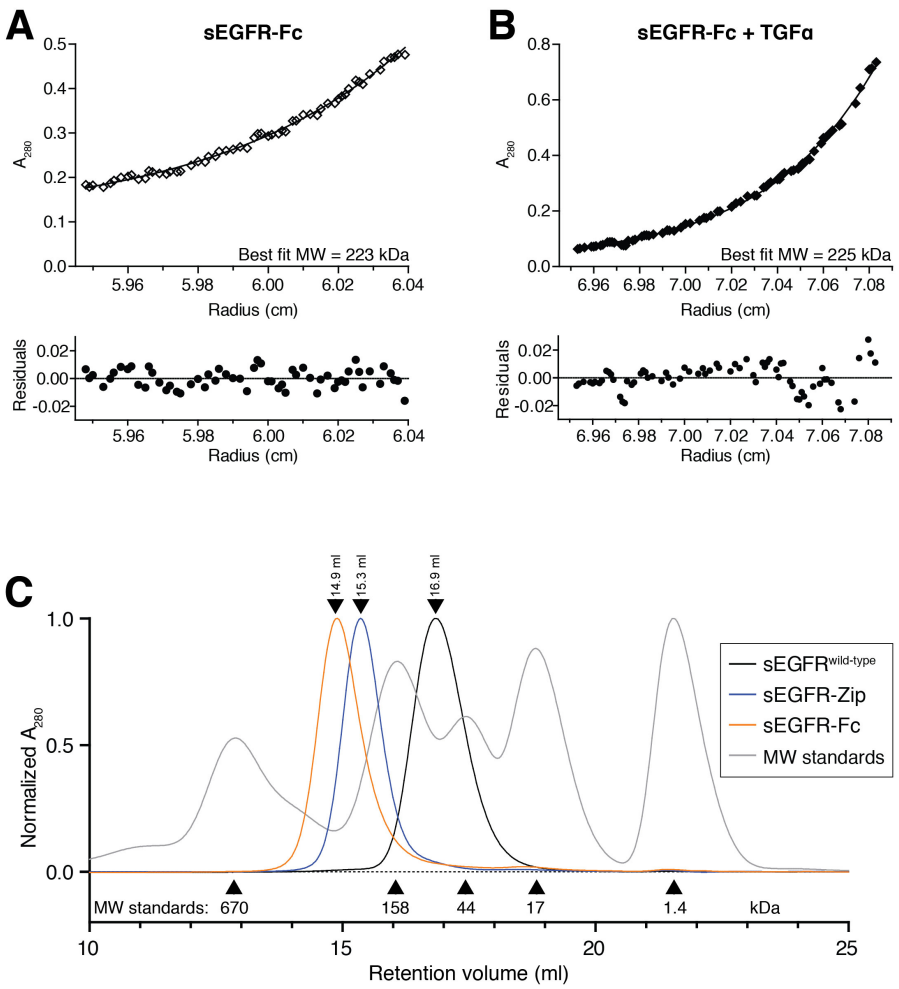


Figure 2.7 Analysis of molecular mass of EGFR-Fc and sEGFR-Zip.

(**A**) Sedimentation equilibrium AUC data, showing single-species fits for data obtained with sEGFR-Fc alone at 10 μM at a speed of 10000 r.p.m. Data were fit using Sedphat (Schuck, 2003) and yielded an excellent fit – with small random residuals – to a single species of 223 kDa. This compares will with the expected mass of an sEGFR-Fc dimer (190 kDa plus ~20% w/w carbohydrate). (**B**) An equivalent experiment to that shown in A for sEGFR-Fc, but with the addition of a 1.2-fold molar excess of TGFα. The best-fit molecular mass is unchanged (at 225 kDa), showing that the dimeric sEGFR-Fc is not further oligomerized upon ligand binding. (**C**) Size exclusion chromatography analysis of sEGFR$^{wild\text{-}type}$ (black curve), s-

EGFR-Zip (blue curve), and sEGFR-Fc (orange curve), injected onto a Superose 6 column at a concentration of 10 μM. The receptor elution volumes reveal that sEGFR-Fc and sEGFR-Zip are both dimers, while $sEGFR^{wild\text{-}type}$ is monomeric as expected. This was performed by Nicholas Bessman.

Table 2.1. ITC data for EGF binding to sEGFR variants.

Values are the mean ± standard deviation of at least three independent experiments. ΔG values are calculated from the mean K_D, and $T\Delta S$ values are obtained by subtracting ΔG from the mean value for ΔH. These data were generated and analyzed by Nicholas Bessman.

sEGFR variant	K_D **(nM)**	ΔH **(kcal/mol)**	ΔG **(kcal/mol)**	$T\Delta S$ **(kcal/mol)**
$sEGFR^{wild\text{-}type}$	78 ± 14/39 ± 4	+6.9 ± 0.5	-9.7 (ITC) /10.1 (FA)	16.6 (ITC)/17.0 (FA)
$sEGFR^{Y251A/R285S}$	74 ± 10	+8.9 ± 1.0	-9.7	18.6
$sEGFR\text{-}Fc^{wild\text{-}type}$	7.8 ± 3	+10.3 ± 0.5	-11.1	21.4
sEGFR-Zipper	8.9 ± 1.1	+12.3 ± 0.6	-11.0	23.3

Table 2.2, related to Table 2.1. ITC data for TGFα binding to sEGFR variants.

Values are the mean ± standard deviation of at least three independent experiments. ΔG values are calculated from the mean K_D, and $T\Delta S$ values are obtained by subtracting ΔG from the mean value for ΔH. These data were generated and analyzed by Nicholas Bessman.

sEGFR variant	K_D **(nM)**	ΔH **(kcal/mol)**	ΔG **(kcal/mol)**	$T\Delta S$ **(kcal/mol)**
$sEGFR^{wild\text{-}type}$	80 ± 16	+8.2 ± 0.8	-9.7	17.9
$sEGFR^{Y251A/R285S}$	82 ± 10	+10.1 ± 1.2	-9.7	19.8
$sEGFR\text{-}Fc^{wild\text{-}type}$	nd	+11.0 ± 0.7	nd	nd

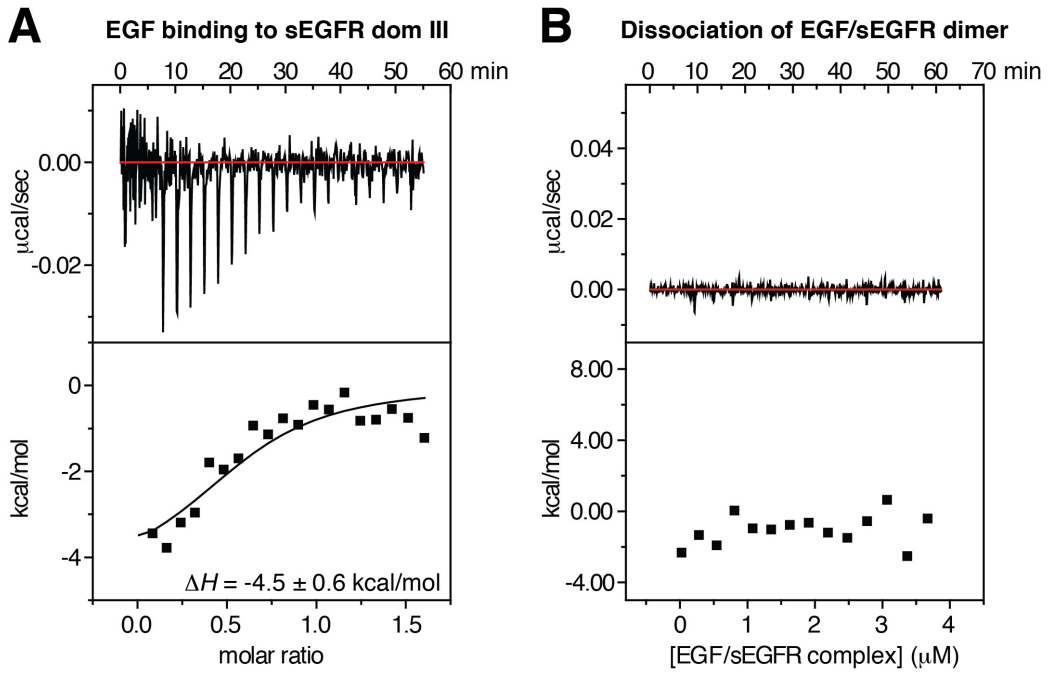


Figure 2.8, ITC of EGF binding to domain III, and EGF/sEGFR dimer dissociation.

(**A**) EGF was titrated into 10 μM sEGFR domain III purified as described (Lemmon, 1997). Fitting to a single binding-site model yielded a K_D value of 3.1 ± 1.9 μM, and a ΔH for EGF binding of -4.5 ± 0.6 kcal/mol. (**B**) To assess the enthalpy associated with dimerization of liganded sEGFR, a dilution experiment was performed in which 17.5 μM sEGFR$^{wild\text{-}type}$ to which a 1.3-fold molar excess of EGF had been added was diluted by injection into an ITC cell containing only buffer. 13 injections of 3 μl each were made, allowing measurement of dissociation heats in the ITC cell over an equilibrium concentration range of ~0.2 – 2.4 μM for the EGF/sEGFR$^{wild\text{-}type}$ complex. Importantly, EGF dissociation from sEGFR$^{wild\text{-}type}$ will be negligible over this concentration range. Because the heat of injection does not change systematically from the first injection to the last, and because the integrated heat of each injection is so low (and can not be distinguished from instrumental noise), we conclude that

the ΔH for dimerization of EGF-bound sEGFR$^{wild\text{-}type}$ must be <<|2| kcal/mol. This work was performed by Nicholas Bessman,

The titrations in Figures 2.2B and 2.2C fit very well to a single entropy-driven (positive ΔH) ligand binding event. In our 1997 studies (Lemmon, 1997), which pre-dated structures of sEGFR, we inferred that this entropy-driven event reflects sEGFR dimerization, and that ligand binding has a small negative ΔH (based on studies of EGF binding to isolated domain III, repeated in **A**). Our new finding that the major entropy-driven event is unaffected by mutations that abolish sEGFR dimerization proves this wrong. Moreover, based on calorimetric dissociation experiments (**B**) we find that dimerization of ligand-bound sEGFR$^{wild\text{-}type}$ has a negligible ΔH. Thus, EGF binding is entropy driven, consistent with another recent report (Alvarenga et al., 2012).

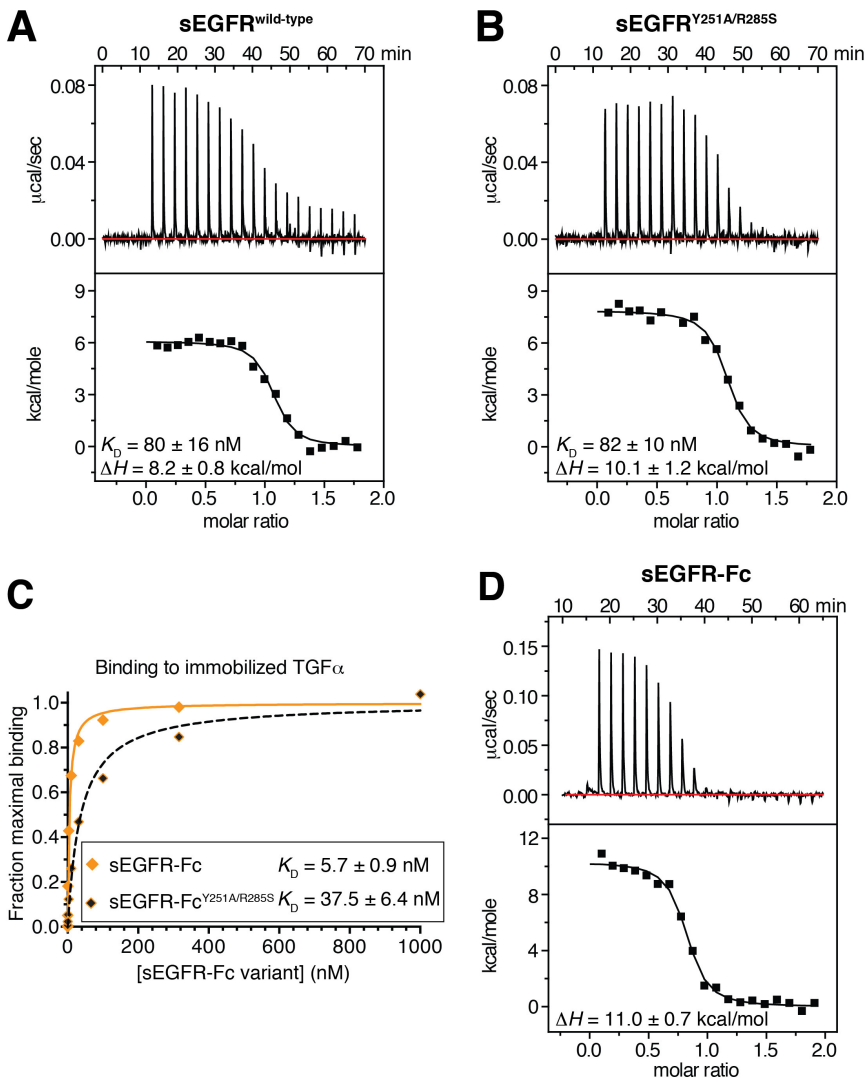


Figure 2.9, related to Figure 2.2. TGFα binding to sEGFR and sEGFR-Fc variants.

(**A**) Representative ITC analysis of TGFα binding to sEGFR$^{wild\text{-}type}$ at 25˚C. TGFα was present in the syringe at 70 μM, and sEGFR$^{wild\text{-}type}$ was present in the calorimeter cell at a concentration of 8 μM. Mean values (± standard deviation) for K_D and ΔH for TGFα binding from three independent experiments are listed in the figure. (**B**) Equivalent experiment to that shown in A for sEGFR$^{Y251A/R285S}$. (**C**) SPR analysis of TGFα binding to sEGFR-Fc with wild-type domain II or with the Y251A/R285S mutations in the domain II dimerization interface. Note that, whereas these mutations had no effect on EGF binding to sEGFR-Fc,

they appear to reduce TGFα binding by ~6-fold, possibly suggesting a slightly different dependence on dimerization for binding of the two ligands – consistent with the observed structural differences in EGF/sEGFR and TGFα/sEGFR complexes (Liu et al., 2012; Wilson et al., 2009). (**D**) ITC analysis of TGFα binding to sEGFR-Fc. TGFα was present in the syringe at 70 μM, and sEGFR-Fc was present in the calorimeter cell at 9 μM. As with EGF binding, TGFα binding to sEGFR-Fc has a higher (positive or unfavorable) enthalpy than binding to $sEGFR^{wild\text{-}type}$ – by 2.8 kcal/mol (compared with 3.4 kcal/mol in the EGF case) – affinity is increased due to entropic effects. This work was performed by Nicholas Bessman.

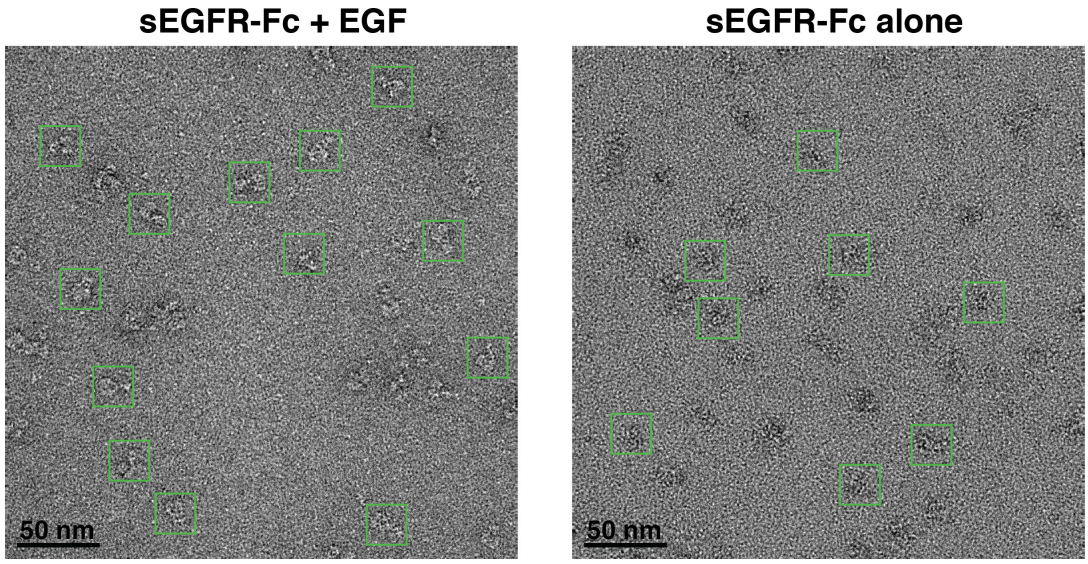


Figure 2.10. Example raw EM images for sEGFR-Fc plus and minus ligand.

Examples of raw image files are shown, from EM studies of negatively stained samples, for sEGFR-Fc plus EGF (left panel) and without ligand (right panel). 150 images such as these were used for single-particle analysis as described in Methods. Example particles are enclosed in green boxes in each representative image. This work was performed and analyzed by Nicholas Bessman.

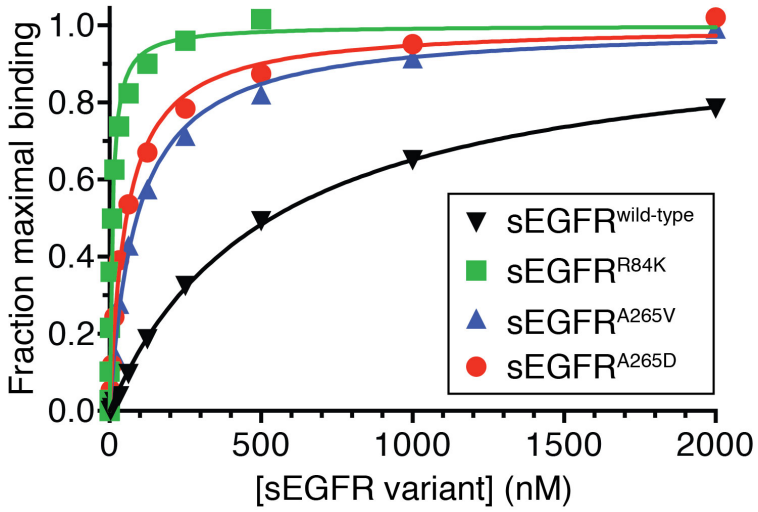


Figure 2.11, related to Figure 2.6. TGFα binding to sEGFR variants harboring glioblastoma mutations.

Binding of each glioblastoma-mutated sEGFR variant to TGFα was analyzed by SPR, using the same approach described for EGF binding in Figure 6B, but with immobilized TGFα instead of EGF. Best fit K_D values for TGFα binding (from three independent experiments) were 532 ± 23 nM ($sEGFR^{wild\text{-}type}$); 8.6 ± 0.9 nM ($sEGFR^{R84K}$); 90 ± 6 nM ($sEGFR^{A265V}$); and 55 ± 4 nM ($sEGFR^{A265D}$). Glioblastoma mutations in EGFR thus enhance TGFα binding to the isolated ECR by 6-62 fold. This work was performed and analyzed by Atrish Bagchi.

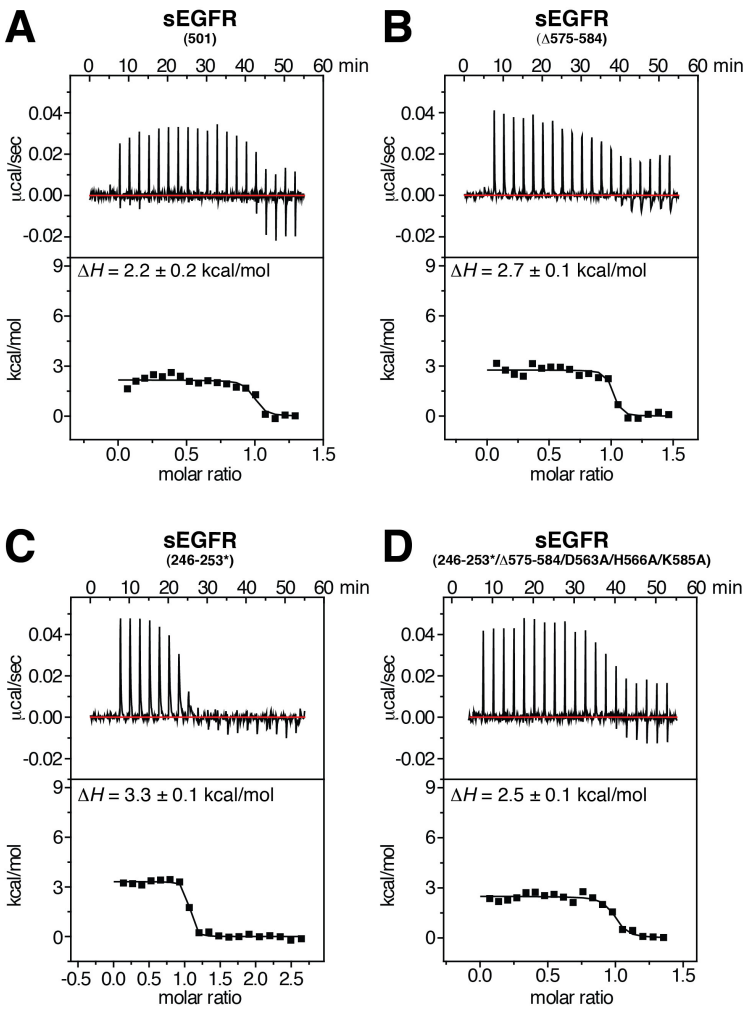


Figure 2.12, related to Figure 2.6. ITC studies of EGF binding to sEGFR with tether-disrupting mutations.

(**A**) EGF was titrated at 25˚ C into a cell containing 10 μM sEGFR(501), a variant of sEGFR that lacks essentially all of domain IV (Elleman et al., 2001), so cannot make the intramolecular tether. K_D for EGF binding to this variant is 7.8 nM (Dawson, 2005), and ΔH is 2.2 ± 0.2 kcal/mol. (**B**) The Δ575-584 mutation deletes a prominent loop from the fifth disulfide-bonded module of domain IV in sEGFR, removing a loop that interacts with the dimerization arm in the tethered sEGFR structure – and thus weakening the tether (Dawson, 2005). K_D for EGF binding to this variant is 32 nM (Dawson, 2005), and ΔH is 2.7 ± 0.1 kcal/mol. (**C**) The 246-253* mutation was initially described by Garrett et al. (2002),

and includes the following 6 mutations, designed to break dimerization arm contacts: Y246E, N247A, T249D, Y251E, Q252A, and M253D. These mutations disrupt the intramolecular tether shown in Figure 1A and the dimerization contacts in Figure 1C. K_D for EGF binding to this variant is 260 nM by SPR (Dawson, 2005), and ΔH is 3.3 ± 0.1 kcal/mol. (**D**) The 246-253*/Δ575-584/D563A/H566A/K585A-mutated variant includes all of the changes in the other tether mutations described here, and also includes mutations of D563, H566, and K585 in domain IV to alanine. These residues contribute directly to the domain II/IV tether. Thus, this variant has all tether contacts as well as the dimerization site mutated. ΔH for EGF binding to this variant is 2.5 ± 0.1 kcal/mol. This was performed by Nicholas Bessman and Atrish Bagchi.

REFERENCES

Alvarenga, M.L., Kikhney, J., Hannewald, J., Metzger, A.U., Steffens, K.-J., Bomke, J., Krah, A., and Wegener, A. (2012). In-depth biophysical analysis of interactions between therapeutic antibodies and the extracellular domain of the epidermal growth factor receptor. Anal. Biochem. *421*, 138-151.

Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., and Svergun, D.I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. J. Appl. Crystallogr. *36*, 1277-1282.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605-1612.

Schuck, P. (2003). On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. Anal. Biochem. *320*, 104-124.

Schuck, P., Perugini, M.A., Gonzales, N.R., Howlett, G.J., and Schubert, D. (2002). Size-distribution analysis of proteins by analytical ultracentrifugation: Strategies and application to model systems. Biophys. J. *82*, 1096-1111.

Shaikh, T.R., Gao, H., Baxter, W.T., Asturias, F.J., Boisset, N., Leith, A., and Frank, J. (2008). SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. Nat. Protocols *3*, 1941-1974

Svergun, D. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J. Appl. Crystallogr. *25*, 495-503.

Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: an extensible image processing suite for electron microscopy. J. Struct. Biol. *157*, 38-46.

Adak, S., Yang, K. S., Macdonald-Obermann, J., and Pike, L. J. (2011). The membrane-proximal intracellular domain of the epidermal growth factor receptor underlies negative cooperativity in ligand binding. J Biol Chem *286*, 45146-45155.

Adams, T. E., Koziolek, E. J., Hoyne, P. H., Bentley, J. D., Lu, L., Lovrecz, G., Ward, C. W., Lee, F. T., Scott, A. M., Nash, A. D.*, et al.* (2009). A truncated soluble epidermal growth factor receptor-Fc fusion ligand trap displays anti-tumour activity in vivo. Growth Factors *27*, 141-154.

Alvarado, D., Klein, D. E., and Lemmon, M. A. (2010). Structural Basis for Negative Cooperativity in Growth Factor Binding to an EGF Receptor. Cell *142*, 568-579.

Alvarado, D., Klein, D.E., Lemmon, M.A. (2009). ErbB2/HER2/Neu resembles an autoinhibited invertebrate EGF receptor. Nature *461*, 287-291.

Alvarenga, M. L., Kikhney, J., Hannewald, J., Metzger, A. U., Steffens, K.-J., Bomke, J., Krah, A., and Wegener, A. (2012). In-depth biophysical analysis of interactions between therapeutic antibodies and the extracellular domain of the epidermal growth factor receptor. Analytical Biochemistry *421*, 138-151.

Arkhipov, A., Shan, Y., Das, R., Endres, N. F., Eastwood, M. P., Wemmer, D. E., Kuriyan, J., and Shaw, D. E. (2013). Architecture and membrane interactions of the EGF receptor. Cell *152*, 557-569.

Bass, J., Kurose, T., Pashmforoush, M., and Steiner, D. F. (1996). Fusion of insulin receptor ectodomains to immunoglobulin constant domains reproduces high-affinity insulin binding in vitro. J Biol Chem *271*, 19367-19375.

Bessman, N. J., and Lemmon, M. A. (2012). Finding the missing links in EGFR. Nat Struct Mol Biol *19*, 1-3.

Brennan, C. W., Verhaak, R. G., McKenna, A., Campos, B., Noushmehr, H., Salama, S. R., Zheng, S., Chakravarty, D., Sanborn, J. Z., Berman, S. H.*, et al.* (2013). The somatic genomic landscape of glioblastoma. Cell *155*, 462-477.

Burgess, A. W., Cho H.S., Eigenbrot C., Ferguson K.M., Garrett T.P., Leahy D.J., Lemmon M.A., Sliwkowski M.X., Ward C.W., Yokoyama S. (2003). An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. Mol Cell *12*, 541-552.

Cantor, C. R., and Schimmel, P. R. (1980). Biophysical Chemistry: Techniques for the study of biological structure and function., (New York, USA: WH Freeman and Co).

Chantry, A. (1995). The kinase domain and membrane localization determine intracellular interactions between epidermal growth factor receptors. J Biol Chem *270*, 3068-3073.

Dawson, J. P., Berger, M. B., Lin, D., Schlessinger, J., Lemmon, M. A., Ferguson, K. M. (2005). EGF receptor dimerization and activation require ligand-induced conformational changes in the dimer interface. Mol Cell Biol *25*, 7734-7742.

Dawson, J. P., Bu, Z., and Lemmon, M. A. (2007). Ligand-Induced Structural Transitions in ErbB Receptor Extracellular Domains. Structure *15*, 942-954.

De Meyts, P. (2008). The insulin receptor: a prototype for dimeric, allosteric membrane receptors? Trends Biochem Sci *33*, 376-384.

De Meyts, P., and Whittaker, J. (2002). Structural biology of insulin and IGF1 receptors: Implications for drug design. Nat Rev Drug Discov *1*, 769-783.

Defize, L. H., Boonstra, J., Meisenhelder, J., Kruijer, W., Tertoolen, L. G., Tilly, B. C., Hunter, T., van Bergen en Henegouwen, P. M., Moolenaar, W. H., and de Laat, S. W. (1989). Signal transduction by epidermal growth factor occurs through the subclass of high affinity receptors. J Cell Biol *109*, 2495-2507.

Elleman, T. C., Domagala, T., McKern, N. M., Nerrie, M., Lonnqvist, B., Adams, T. E., Lewis, J., Lovrecz, G. O., Hoyne, P. A., Richards, K. M.*, et al.* (2001). Identification of a determinant of epidermal growth factor receptor ligand-binding specificity using a truncated, high-affinity form of the ectodomain. Biochemistry *40*, 8930-8939.

Endres, N. F., Das, R., Smith, A. W., Arkhipov, A., Kovacs, E., Huang, Y., Pelton, J. G., Shan, Y., Shaw, D. E., Wemmer, D. E.*, et al.* (2013). Conformational coupling across the plasma membrane in activation of the EGF receptor. Cell *152*, 543-556.

Ferguson, K. M., Berger, M. B., Mendrola, J. M., Cho, H. S., Leahy, D. J., and Lemmon, M. A. (2003). EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. Mol Cell *11*, 507-517.

Ferguson, K. M., Darling, P. J., Mohan, M. J., Macatee, T. L., Lemmon, M. A. (2000). Extracellular domains drive homo- but not hetero-dimerization of erbB receptors. EMBO J *19*, 4632-4643.

Garrett, T. P. J., McKern, N. M., Lou, M., Elleman, T. C., Adams, T. E., Lovrecz, G. O., Zhu, H. J., Walker, F., Frenkel, M. J., Hoyne, P. A., Jorissen, R. N., Nice, E. C., Burgess, A. W., Ward, C. W. (2002). Crystal structure of a truncated epidermal growth factor receptor domain bound to transforming growth factor alpha. Cell *110*, 763-773.

Graus-Porta, D., Beerli, R. R., Daly, J. M., and Hynes, N. E. (1997). ErbB-2, the preferred heterodimerization partner of all ErbB receptors, is a mediator of lateral signaling. EMBO J *16*, 1647-1655.

Hoyne, P. A., Cosgrove, L. J., McKern, N. M., Bentley, J. D., Ivancic, N., Elleman, T. C., and Ward, C. W. (2000). High affinity insulin binding by soluble insulin receptor extracellular domain fused to a leucine zipper. FEBS Letts *479*, 15-18.

Jaiswal, B. S., Kljavin, N. M., Stawiski, E. W., Chan, E., Parikh, C., Durinck, S., Chaudhuri, S., Pujara, K., Guillory, J., Edgar, K. A.*, et al.* (2013). Oncogenic ERBB3 mutations in human cancers. Cancer Cell *23*, 603-617.

Jones, J. T., Akita, R.W., Sliwkowski, M.X. (1999). Binding specificities and affinities of egf domains for ErbB receptors. FEBS Letters *447*, 227-231.

Jura, N., Endres, N.F., Engel, K., Deindl, S., Das, R., Lamers, M.H., Wemmer, D.E., Zhang, X., Kuriyan, J. (2009). Mechanism for activation of the EGF receptor catalytic domain by the juxtamembrane segment. Cell *137*, 1293-1307.

Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. J Appl Crystallogr *36*, 1277-1282.

Lax, I., Mitra, A. K., Ravera, C., Hurwitz, D. R., Rubinstein, M., Ullrich, A., Stroud, R. M., and Schlessinger, J. (1991). Epidermal growth factor (EGF) induces oligomerization of soluble, extracellular, ligand-binding domain of EGF receptor. A low resolution projection structure of the ligand-binding domain. J Biol Chem *266*, 13828-13833.

Lee, J. C., Vivanco, I., Beroukhim, R., Huang, J. H., Feng, W. L., DeBiasi, R. M., Yoshimoto, K., King, J. C., Nghiemphu, P., Yuza, Y.*, et al.* (2006). Epidermal growth factor receptor activation in glioblastoma through novel missense mutations in the extracellular domain. PLoS Med *3*, e485.

Lemmon, M. A. (2009). Ligand-induced ErbB receptor dimerization. Exp Cell Res *315*, 638-648.

Lemmon, M. A., Bu, Z., Ladbury, J. E., Zhou, M., Pinchasi, D., Lax, I., Engelman, D. M., Schlessinger, J. (1997). Two EGF molecules contribute additively to stabilization of the EGFR dimer. EMBO J *16*, 281-294.

Lemmon, M. A., Schlessinger, J., and Ferguson, K. M. (2014). The EGFR family: not so prototypical receptor tyrosine kinases. Cold Spring Harb Perspect Biol *6*, a020768.

Li, S., Schmitz, K. R., Jeffrey, P. D., Wiltzius, J. J., Kussie, P., and Ferguson, K. M. (2005). Structural basis for inhibition of the epidermal growth factor receptor by cetuximab. Cancer Cell *7*, 301-311.

Li, Y., Macdonald-Obermann, J., Westfall, C., Piwnica-Worms, D., and Pike, L. J. (2012). Quantitation of the effect of ErbB2 on epidermal growth factor receptor binding and dimerization. J Biol Chem *287*, 31116-31125.

Liu, P., Cleveland, T. E. t., Bouyain, S., Byrne, P. O., Longo, P. A., and Leahy, D. J. (2012). A single ligand is sufficient to activate EGFR dimers. Proc Natl Acad Sci U S A *109*, 10861-10866.

Livingstone, J. R., Spolar, R. S., and Record, M. T., Jr. (1991). Contribution to the thermodynamics of protein folding from the reduction in water-accessible nonpolar surface area. Biochemistry *30*, 4237-4244.

Lu, C., Mi, L. Z., Grey, M. J., Zhu, J., Graef, E., Yokoyama, S., and Springer, T. A. (2010). Structural evidence for loose linkage between ligand binding and kinase activation in the epidermal growth factor receptor. Mol Cell Biol *30*, 5432-5443.

Lu, C., Mi, L. Z., Schurpf, T., Walz, T., and Springer, T. A. (2012). Mechanisms for kinase-mediated dimerization of the epidermal growth factor receptor. J Biol Chem *287*, 38244-38253.

MacDonald, J. L., Pike, L.J. (2008). Heterogeneity in EGF-binding affinities arises from negative cooperativity in an aggregating system. Proc Natl Acad Sci U S A *105*, 112-117.

Macdonald-Obermann, J. L., and Pike, L. J. (2009). The intracellular juxtamembrane domain of the epidermal growth factor (EGF) receptor is responsible for the allosteric regulation of EGF binding. J Biol Chem *284*, 13570-13576.

Magun, B. E., Matrisian, L. M., and Bowden, G. T. (1980). Epidermal growth factor. Ability of tumor promoter to alter its degradation, receptor affinity and receptor number. J Biol Chem *255*, 6373-6381.

Mattoon, D., Klein, P., Lemmon, M. A., Lax, I., and Schlessinger, J. (2004). The tethered configuration of the EGF receptor extracellular domain exerts only a limited control of receptor function. Proc Natl Acad Sci USA *101*, 923-928.

Mi, L. Z., Lu, C., Li, Z., Nishida, N., Walz, T., and Springer, T. A. (2011). Simultaneous visualization of the extracellular and cytoplasmic domains of the epidermal growth factor receptor. Nat Struct Mol Biol *18*, 984-989.

Ogiso, H., Ishitani, R., Nureki, O., Fukai, S., Yamanaka, M., Kim, J. H., Saito, K., Sakamoto, A., Inoue, M., Shirouzu, M., Yokoyama, S. (2002). Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. Cell *110*, 775-787.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem *25*, 1605-1612.

Red Brewer, M., Choi, S. H., Alvarado, D., Moravcevic, K., Pozzi, A., Lemmon, M. A., and Carpenter, G. (2009). The juxtamembrane region of the EGF receptor functions as an activation domain. Mol Cell *34*, 641-651.

Saphire, E. O., Parren, P. W., Pantophlet, R., Zwick, M. B., Morris, G. M., Rudd, P. M., Dwek, R. A., Stanfield, R. L., Burton, D. R., and Wilson, I. A. (2001). Crystal structure of a neutralizing human IGG against HIV-1: a template for vaccine design. Science *293*, 1155-1159.

Schlessinger, J. (1986). Allosteric regulation of the epidermal growth factor receptor kinase. J Cell Biol *103*, 2067-2072.

Schuck, P. (2003). On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. Anal Biochem *320*, 104-124.

Schuck, P., Perugini, M. A., Gonzales, N. R., Howlett, G. J., and Schubert, D. (2002). Size-distribution analysis of proteins by analytical ultracentrifugation: Strategies and application to model systems. Biophys J *82*, 1096-1111.

Shaikh, T. R., Gao, H., Baxter, W. T., Asturias, F. J., Boisset, N., Leith, A., and Frank, J. (2008). SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. Nat Protocols *3*, 1941-1974

Surinya, K. H., Forbes, B. E., Occhiodoro, F., Booker, G. W., Francis, G. L., Siddle, K., Wallace, J. C., and Cosgrove, L. J. (2008). An investigation of the ligand binding properties and negative cooperativity of soluble insulin-like growth factor receptors. J Biol Chem *283*, 5355-5363.

Svergun, D. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J Appl Crystallogr *25*, 495-503.

Sweeney, C., and Carraway, K. L., 3rd. (2000). Ligand discrimination by ErbB receptors: differential signaling through differential phosphorylation site usage. Oncogene *19*, 5568-5573.

Tang, G., Peng, L., Baldwin, P. R., Mann, D. S., Jiang, W., Rees, I., and Ludtke, S. J. (2007). EMAN2: an extensible image processing suite for electron microscopy. J Struct Biol *157*, 38-46.

Vivanco, I., Robins, H. I., Rohle, D., Campos, C., Grommes, C., Nghiemphu, P. L., Kubek, S., Oldrini, B., Chheda, M. G., Yannuzzi, N.*, et al.* (2012). Differential sensitivity of glioma-versus lung cancer-specific EGFR mutations to EGFR kinase inhibitors. Cancer Disc *2*, 458-471.

Walker, F., Orchard, S. G., Jorissen, R. N., Hall, N. E., Zhang, H. H., Hoyne, P. A., Adams, T. E., Johns, T. G., Ward, C., Garrett, T. P. J.*, et al.* (2004). CR1/CR2 interactions modulate the functions of the cell surface epidermal growth factor receptor. J Biol Chem *279*, 22387-22398.

Wilson, K. J., Gilmore, J. L., Foley, J., Lemmon, M. A., and Riese, D. J., 2nd (2009). Functional selectivity of EGF family peptide growth factors: implications for cancer. Pharmacol Ther *122*, 1-8.

CHAPTER THREE

Towards understanding allosteric regulation of intact EGFR using hydrogen/deuterium exchange coupled to mass spectrometry

SUMMARY

The epidermal growth factor receptor (EGFR) was the first growth factor receptor for which ligand-induced dimerization was shown to be a relevant mechanism for receptor activation. Recent studies, however, have shown that ligand-independent inactive dimers also exist at the cell surface in both human EGFR and EGFR orthologues in *Drosophila melanogaster* and *Caenorhabditis elegans* (Freed, 2015). In chapter two, we discovered that ligand binding and dimerization are competitive processes, and alterations in the thermodynamic linkage of these processes can contribute to EGFR oncogenic dysregulation (Bessman, 2014). We hypothesized that allosteric regulation of EGFR by the GBM mutations studied in Chapter Two may alter the relative energetic contributions of extra and intra-cellular interactions to EGFR dimerization, and proposed to study the nature of this dysregulation by hydrogen/deuterium exchange coupled to mass spectrometry analysis of intact EGFR in detergent micelles and unilamellar proteoliposomes. In this chapter, I discuss progress towards application of this approach to study allosteric regulation of EGFR.

INTRODUCTION

In Chapter Two, we identified the EGFR domain I/II interface as a novel site of autoinhibitory interactions (Bessman, 2014). We also showed that glioblastoma-linked mutations in this interface do not simply promote ligand-independent EGFR activation by enhancing domain II-mediated dimerization in the absence of contributions from EGFR intracellular domains and the plasma membrane. Surprisingly, these activating glioblastoma mutations in the extracellular region of EGFR sensitize the receptor to inhibition by lapatinib (Vivanco, 2012), a kinase inhibitor that, based upon crystallographic data, binds preferentially to the kinase inactive conformation and also has a relatively slow off-rate (Wood, 2004). Taken together with the observations that there are inactive, preformed dimers at the cell surface (Yu, 2002) and the extracellular regions of the *Drosophila melanogaster* (Alvarado, 2009) and *C. elegans* orthologs of EGFR dimerize in the absence of ligand (Freed, 2015), we hypothesized that EGFR is regulated by long-range, allosteric conformational changes. Taken together, it is increasingly clear that the role of the EGFR extracellular region in mediating receptor activation may involve long range allosteric effects in response to ligand binding in addition to its role in mediating receptor dimerization, and that there are autoinhibitory mechanisms within receptor dimers.

Although there are extensive structural and biochemical data on isolated individual domains of the EGF receptor (Dawson, 2005; Dawson, 2007; Ferguson, 2003; Red Brewer, 2009; Zhang, 2006), there are only a handful of studies that have biophysically assessed the coupling of extracellular and intracellular conformations (Lu, 2010; Lu, 2012; Mi, 2011; Wang, 2011). Some of these studies directly point to an interdependence between extracellular and intracellular conformations. For example, kinase inhibitors that preferentially bind to the active kinase configuration also may enhance extracellular dimerization (Lu, 2012). In

addition, there has been a suggestion that modulation of the intracellular juxtamembrane region affects affinity (Macdonald-Obermann, 2009) and cooperativity (Adak, 2011) of ligand binding. The hypothesis of conformational 'cross talk' and 'inside-out' signaling mechanisms across the membrane and its importance in EGF receptor activation has also been challenged by a study that argued that conformational linkage of the extracellular and intracellular domains is weak (Lu, 2010).

These analyses demonstrate that we do not yet understand the relationship between extracellular and intracellular conformation of EGFR. I proposed to study the relationship between ligand-induced extracellular conformational changes and intracellular tyrosine using hydrogen/deuterium exchange coupled to mass spectrometry (HDX/MS). Backbone amide hydrogens exchange with solvent hydrogens at rates that are sensitive to both secondary structure and internal motions. As amide hydrogens can only exchange with solvent hydrogens (or deuteriums) when backbone hydrogen bonding in secondary structures are transiently broken, the rates of amide hydrogen exchange are sensitive to backbone dynamics and structural fluctuations (Skinner, 2012). The kinetics of amide hydrogen exchange can be followed on a residue level by NMR or on a peptide level by mass spectrometry (Skinner, 2012), although residue resolution can now be achieved by analysis of overlapping peptides and isotopic envelope shape information (Kan, 2013). These data can give valuable information about changes in secondary structure and dynamics given a perturbation, and therefore it may serve as a useful structural tool to analyze the oncogenic EGFR mutations discussed in Chapter 2.

In this chapter, I discuss preliminary efforts to assess the long range conformational effects of ligand binding to the extracellular region in full-length receptors by hydrogen/deuterium exchange coupled to mass spectrometry. I report EGF-dependent protection factors in the EGFR extracellular region. I also made substantial progress toward application of this technique to studying full-length EGFR in detergent micelles. Future

studies focused on such analyses of EGFR reconstituted in a physiologically relevant setting may gain insight into the mechanisms underlying receptor oncogenic dysregulation and lipid-mediated autoinhibition.

RESULTS AND DISCUSSION

Hydrogen/Deuterium exchange analysis of the isolated EGFR extracellular region in the presence and absence of EGF

We performed hydrogen/deuterium exchange coupled to mass spectrometry experiments on the recombinantly produced and purified EGFR extracellular region (mature amino acids 1-618, sEGFR) in the presence and absence of saturating concentrations of EGF. We observed EGF-driven hydrogen exchange protection in regions of sEGFR known to be involved in direct contacts with EGF, and dimerization (Figure 3.1). Briefly, peptides spanning the N-terminal region of domain I, and the β strands in domain III that make direct contacts with EGF undergo HDX protection. The magnitude of HDX protection in EGF driven protection ranges from ~10,000 fold in some peptides in a random coil part of domain I to between ~10-100 fold in peptides in the β helix solenoid of EGFR domain III and the dimerization interface in domain II. In addition, we observed ~100-fold allosteric HDX protection in the C-terminal region of domain III, at the domain III/IV interface. This protection could be a result of the stabilization of the domain III/IV boundary in a ligated EGFR dimer. There is also a very small (<10-fold) deprotection effect on a loop in domain III (aa. 370-380). For instance, peptide 369-381 +3 exhibits a 4 % increase in HX at the 24,000 second timepoint with excellent statistical significance ($p=0.0005$) by an unpaired t-test. Other peptides in this region require more independent replicates for measure of statistical significance.

This HDX protection likely results from the secondary structure stabilization at protein/protein interfaces, and, in the context of ligand-binding peptides in domain I, from structure acquisition. The relative magnitude of these changes shown in Figure 3.1 suggest an energetic distribution of the relative contributions of stabilization that result from high affinity EGF binding to domains I and III (K_d ~150 nM) and dimerization (K_d ~1 μM). These results corroborate the extant view that HDX protection factors can be useful tools to directly quantitatively analyze the relative stabilization of secondary structure at protein/protein interfaces (Skinner, 2012). The deprotection effect on the aa. 370-380 loop may result from the loss of interactions with domain II in this region of domain III. In principle, one could conceivably use this approach to also analyze receptor/detergent micelle, receptor/lipid, and receptor/downstream effector interactions and secondary structure and dynamics, as has been reported recently for G-protein coupled receptors (Chung, 2011; Duc, 2015; Shukla, 2014).

My preliminary HDX analysis of the EGFR extracellular region is an important step towards eventual application of this technique and other complementary structural mass-spectrometry based techniques, like fast photochemical oxidation coupled to mass spectrometry (FPOP), which measures changes in solvent accessibility (Chen, 2010), to studying the allosteric regulation of full-length EGFR. These mass spectrometry-based structural techniques can be applied to systems that are not currently amenable to X-ray crystallography, such as large complexes or membrane proteins.

In performing these experiments, I encountered two main technical challenges imposed by both of the types of subdomains in the EGFR extracellular region. Domains II and IV in the EGFR extracellular region are cysteine rich domains, and both are structurally stabilized by the presence of the vast majority of EGFR's disulfide bonds. Domains I and III are the ligand binding domains, and each contain a β helix solenoid structure.

The first technical challenge was MS/MS coverage of peptides in the cysteine rich domains. Under the conditions (pH~ 2.5-3) that minimize back exchange for the HDX experiment, the reduction potential of common reducing agents such as tris(2-carboxyethyl)phosphine (TCEP) is mitigated, effectively reducing the efficiency of disulfide reduction, and thereby lowering the intensities of both MS1 and MS2 spectra of the reduced species. Because the acid proteases commonly used for HDX (such as pepsin) are relatively nonspecific, computational identification of disulfide crosslinked peptides remains challenging. We attempted to use empirically observed cut sites from MS/MS analysis as a restraint to identify disulfide-linked peptides, but this approach did not significantly improve coverage in EGFR domains II and IV (Z.-Y. Kan, S. W. Englander, personal communication). As an alternate approach, four iterative rounds of MS/MS using exclusion lists and inclusion lists were used to identify a comprehensive list of identified peptides in non-deuterated samples of the EGFR extracellular region. This approach yielded a library of 392 peptides of sEGFR, with ~85% coverage of the extracellular region (example coverage of one of these MS/MS runs is shown in Figure 3.2).

While MS/MS coverage of nondeuterated sEGFR in its cysteine rich domains was significantly improved through use of a rapid dilution approach (detailed in Methods), significant loss of coverage in low abundant species in cysteine rich domains II and IV occurred in analysis of MS1 spectra of deuterated samples. Similar loss of coverage has been previously reported (Bobst, 2014; Burke, 2008), and could result from both reverse phase chromatographic broading of the deuterated peptides (we used a rather large retention time window for this reason, but this did not help). In addition, statistical validation of the isotopic envelopes implemented in ExMS relies on high spectral resolution and signal to noise (Kan, 2011). Such data quality may be lost for low abundance peptides in the cysteine rich domains of EGFR. We did manage to observe HDX protection for some peptides with cysteines in domain II, such as one in the ‘buttressing’ interaction in domain II (Dawson,

2005) (aa. 275-283, Fig. 3.1), suggesting the optimized rapid dilution approach is useful. However, this approach comes with a greater cost in back exchange, averaging ~25%, than can potentially be achieved in the cooled box reverse phase HPLC setup I used (Walters, 2012). Even greater back exchange has been previously observed for offline protease digestion approaches developed to mitigate the MS/MS coverage problem in disulfide rich regions of proteins (Burke, 2008), highlighting the importance of developing alternate approaches to simultaneously minimize back exchange while maximizing reduction efficiency of disulfides for mass spectrometry identification.

An alternative approach that has been recently described uses electron transfer or capture dissociation (ETD/ECD) based fragmentation to directly cleave the thiol-thiol bond in the gas phase (Bobst, 2014). While hydrogen/deuterium scrambling has previously been observed with this type of fragmentation, gaining interpretable information about EGFR and other receptors with HDX analysis does not necessarily require single residue resolution as has recently been described for smaller globular systems (Kan, 2013). ETD/ECD fragmentation has shown dramatic improvement in coverage for another disulfide-rich glycoprotein transferrin, which is about the same size as the EGFR extracellular region (Bobst, 2014), and is therefore a promising future direction for HDX/MS analysis of disulfide rich proteins like EGFR.

The second technical challenge was to prepare a fully deuterated control sample to calculate peptide-resolved back exchange correction factors. Through screening many conditions, the closest conditions for full deuteration of sEGFR were incubation of sEGFR in 0.08% FA 150 mM NaCl pH 2.4 at room temperature for 65 hours and at 70 degrees for 5 minutes. The distribution of fractional deuteration of peptides in sEGFR under these conditions is compared to a positive control (cytochrome C, for which full deuteration conditions are widely established), in Figure 3.3 (Hu, 2016). The midpoint of the two distributions are about the same, suggesting that the majority of the peptides are indeed fully

deuterated. However, sEGFR contains a 'tail' of peptides that are still not fully deuterated under these conditions (as shown in Figure 3.3A). These peptides are located within the β helix solenoid structures of domains I and III. In addition, aggregation of sEGFR under these conditions decreases the effective concentration injected into the mass spectrometer and reduces coverage in domains II and IV. While the scale factors for peptide specific back exchange are not necessary for comparative HDX analyses of the same peptides, they do limit the analyses to exclude quantitative comparison of distinct peptides, and therefore the raw amide hydrogen exchange rates. This would preclude any analysis of sEGFR dynamics, and would limit any analysis to changes in sEGFR dynamics by a perturbation (i.e. mutations and ligand binding).

Progress towards purification and application of HDX to analysis of full length EGFR

I collaborated with the laboratory of Daniel Leahy (Johns Hopkins), which has previously described an approach to reconstitute detergent solubilized EGFR lacking the C-terminal tail that is phosphorylated upon EGF-induced activation (amino acids 1-998), or tEGFR (Qiu, 2009; Wang, 2011; Wang, 2013), to analyze MS/MS coverage of near-full length tEGFR construct. The tEGFR receptor was purified with an EGF-competitive mAb and eluted with EGF, and supplied in a dodecylmaltoside-containing buffer. We performed preliminary MS/MS coverage analysis of this receptor construct, using a similar dilution approach described above for sEGFR (Figure 3.2B). We were able to identify peptides spanning the entire receptor construct, in both extra- and intra-cellular regions of tEGFR (Figure 3.2B). We did observe many more peptides in disulfide rich regions in the extracellular region of EGFR than coverage experiments for sEGFR (Figure 3.2A), presumably due to the contribution of the reducing agent DTT in the buffer and incubation at room temperature for 10 minutes. We did not obtain peptides in the first ten residues of the intracellular juxtamembrane region,

perhaps due to inaccessibility of the transmembrane (TM) domains to pepsin, producing long peptides too hydrophobic to be eluted from a C18 column under these conditions.

As the analysis possible with tEGFR is inherenty limited by the mAb-based purification approach (it can only be purified in complex with an Fab of an EGF-competitive antibody, such as cetuximab or EGF), we decided to design an alternate approach to purify full-length EGFR. The baculovirus expression system was used to produce recombinant EGFR in Sf9, HiFive, and TnI insect cells. Full length EGFR expressed the best in Sf9 cells, with an optimal harvest time of 48-55 hours post infection. Receptor proteins with a C-terminal hexahistidine tag were purified from the isolated membranes of Sf9 cells by TALON purification followed by anion exchange and size exclusion chromatography. Full-length EGFR (FL-EGFR) solubilized in CHAPS eluted primarily as a presumably monomeric peak by size exclusion chromatography (MW ~170 kDa), with a 'shoulder' indicating possible aggregation. These purifications were done within 30 hours of lysis to limit time-dependent loss of tyrosine kinase activity of the detergent-solubilized receptors, with a yield of ~20 μg/L Sf9 cells. I also made Flag- and Strep-tagged version of full-length EGFR, and the purification efficiency of EGFR containing either of these tags is higher than that of a hexahistidine tagged EGFR.

Preliminary autophosphorylation assays were also performed with CHAPS-solubilized full-length EGFR (detailed in Methods). EGFR autophosphorylation experiments were performed with protein after ion exchange purification (prior to size exclusion chromatography) to evaluate if the tyrosine kinase module would be catalytically active once subject to purification and to mitigate concerns regarding time-dependent loss of kinase activity. Protein >70% pure by Coomassie staining (not shown) at this stage of purification, and was concentrated to an OD_{280} of 0.3 (~2 μM) for autophosphorylation experiments. Autophosphorylation was initiated by addition of 1 mM ATP (pH 7) and 10 mM $MgCl_2$ at room temperature for 10 minutes with or without 40 μM EGF (detailed in Methods). Reactions

were quenched by addition of reducing 3x Laemmli sample buffer, and analyzed by immunoblotting using anti-PY20 and a phosphospecific anti-pY1068 antibody (Santa Cruz) and anti-EGFR murine antibody cocktail Ab-12 (Thermo Scientific) as shown in Figure 3.6. There is some spectral overlap between the anti-pY1068 and anti-EGFR antibodies in the immunoblot shown in Figure 3.6A (evident by the yellow color in lanes 5 and 8, but analysis of other lanes shows approximate equal loading of EGFR). Nevertheless, as shown in figure 3.6A, EGFR autophosphorylation is dependent on addition of both ATP and $MgCl_2$, suggesting that the tyrosine kinase module is active under these conditions. However, addition of EGF under these conditions does not measurably increase the tyrosine autophosphorylation of EGFR in CHAPS. These results are consistent with the first detergent-solubilized preparations of EGFR which had noted an absence of EGF-dependent increase in autophosphorylation unless a preincubation step was included (Cohen, 1980), which I speculate could be due to an increase in the kinetic activation barrier of dimerization imposed by detergent micelles. An alternate possibility is that the protein is in aggregates, consistent with the subsequent size exclusion analysis of these preparations showing an 'aggregated' shoulder (Figure 3.5A), and aggregation causes an increase in ligand-independent tyrosine kinase domain activation (Yarden, 1987). A third possibility is that there is a loss of autoinhibition due to loss of interactions with lipids, and because of the CHAPS micelle (although EGF stimulation has been previously observed of tEGFR in dodecylmaltoside micelles) (Wang, 2011; Wang, 2013). Protein is >80% pure by Coomassie staining after size exclusion chromatography (Figure 3.5B).

Overall, my progress towards application of HDX/MS to evaluating long-range allosteric regulation in EGFR is still at a very preliminary stage. However, HDX/MS, coupled to other MS-based structural techniques, has the potential to reveal modes of oncogenic allosteric dysregulation of EGFR, which may rationalize the differences in lapatinib sensitivity

of the extracellular GBM mutations studied in Chapter 2 (Bessman, 2014), as well as structural mechanisms for ligand discrimination in EGFR (Ronan, 2016).

METHODS

Protein purification and constructs.

sEGFR constructs were purified as described previously (Ferguson, 2000). EGF was purchased from Chemicon, inc. Full length EGFR constructs for purification were made by amplifying EGFR with a 5' KpnI site and a 3' Xho site, double digesting the amplified DNA, and ligating a gel purified amplified product into a digested pfastBac vector. The expression tag was inserted at the amplification step. All constructs were verified by DNA sequencing.

Full-length EGFR was purified from Sf9 cells using a baculovirus system. Sf9 cells were infected with 10 ml/L of P2 baculovirus, and harvested after 50 hours. Cell pellets were washed in ice-cold PBS, and lysed in an isotonic buffer conditions – 25 mM HEPES pH 7.5 200 mM NaCl, 15 % glycerol, 1x mammalian protease inhibitors (Roche, Palo Alto, CA) 1 mM PMSF, using sonication (though a pressure-based French press system would likely be preferable) on ice. Cell debris was removed by centrifugation at 10000g (a range between 500g and 30000g was tried and 10000g gave the best purification with least loss, data not shown) for 10 minutes at 4 degrees. Membranes were isolated from the supernatant by ultracentrifugation of the resulting lysate at 150,000g for 1.5 hours. Membranes were resuspended in 25 mM HEPES pH 8 200 mM NaCl, 15 % glycerol, 1x mammalian protease inhibitors (Roche, Palo Alto, CA) 1 mM PMSF with a dounce homogenizer. Detergent extraction was performed by direct addition of the shown concentrations of the relevant detergents and nutating at 4 degrees for 1.5 hours (Figure 3.4). CHAPS and DDM were

taken forward, as a previous purifications for EGFR and tEGFR have been described using these detergents (Coskun, 2011; Qiu, 2009). Unsolubilized material was removed by ultracentrifugation at 100,000g for 1 hour. The resulting supernatant was incubated by batch binding to 0.25 ml of Talon (Co^{2+} affinity) resin/L Sf9 cells (pre-equilbrated with 2 column volumes of the same conditions as the lysate) for 1 hour, and washed with four column volumes of 25 mM HEPES 5 mM imidazole 1% CHAPS 1x mammalian protease inhibitor cocktail (Roche, Palo Alto, CA), 15% glycerol, 1 mM PMSF, 200 mM NaCl, pH 8, and eluted in the same buffer containing 50 mM imidazole. Protein was further purified by anion exchange chromatography using a resource Q column (GE Healthcare). Purified EGFR fractions in CHAPS were pooled and further purified by size exclusion chromatography with a Superose 6 column (GE Healthcare). Full length-EGFR in CHAPS eluted by size exclusion chromatography predominantly as two species – the major of which is likely monomeric with an aggregated or oligomeric 'shoulder'. The identitiy of this protein was further confirmed by tandem MS/MS analysis (Wistar Proteomic Facility, data not shown). Similar purification results were also observed for full-length EGFR purified in dodecylmaltosize (DDM), and purified receptor in 25 mM HEPES 150 mM NaCl 0.05% DDM 1x mammalian protease inhibitor cocktail (Roche, Palo Alto, CA) 15 % glycerol 1 mM PMSF pH 8. This protein was confirmed to bind to immobilized EGF in a Biacore binding assay under these conditions (data not shown).

Near full length EGFR (truncated at amino acid 998), or tEGFR, bound to EGF was purified from HEK293 GnTi- cells as previously described (Qiu, 2009; Wang, 2011; Wang, 2013), and was generously provided by Lily Raines and Daniel Leahy of Johns Hopkins University School of Medicine.

Autophosphorylation reactions.

The autophosphorylation activity of CHAPS-solubilized full-length EGFR was performed after resource Q fraction. Protein was in 25 mM HEPES ~200 mM NaCl 1% CHAPS 15% glycerol 1x mammalian protease inhibitor cocktail pH 8.0 (Roche, Palo Alto, CA) 1 mM PMSF and was concentrated to a buffer reference subtracted Abs_{280} of 0.3 using a 100 kDa cutoff concentrator at 500g (at 4 degrees). EGFR Buffer for the experiment was 25 mM HEPES 150 mM NaCl 1% CHAPS 15% glycerol 1x mammalian protease inhibitor cocktail pH 8.0 (Roche, Palo Alto, CA) 1 mM PMSF. Stocks were 0.5 M $MgCl_2$, 50 mM ATP (pH 7), and 400 μM EGF (in 50 mM HEPES 150 mM NaCl pH 8.0) – all sterile-filtered. Reactions were initiated by mixing 7.6 μl EGFR in indicated conditions with 2.4 μl EGFR buffer, or 1.6 μl EGFR buffer with 0.8 μl of ATP alone, $MgCl_2$ alone, or EGF alone, or 0.8 μl EGFR buffer with combinations of the other two, or 0.8 μl of ATP with 0.8 μl MgCl2 and 0.8 μl EGF.

H/DX reactions

Deuterium on-exchange for EGFR extracellular region was carried out at room temperature by adding 1.2 μl of sEGFR stock (125 μM) in 50 mM HEPES pH 8 150 mM NaCl to 0.3 μl of this same buffer or EGF resuspended in this buffer (at a stock concentration of 625 μM). Reactions were initiated by addition of 6 μl of deuterium on-exchange buffer (150 mM NaCl in D_2O) so that the final reaction conditions 80% D_2O 10 mM HEPES pH 8 (pD~8.4) 150 mM NaCl. Reactions were quenched with addition of 20 μl of pre-chilled QB (0.1 % FA pH 2.8 6.2 M Guanidine 500 mM TCEP) for 30 sec. then diluted with 34.5 μl FB (0.8% FA 10% glycerol) to make final 180 mM TCEP 2 M Gndn-HCl and final volume for each reaction 60 μl. These samples were then immediately frozen in liquid nitrogen. The samples were stored at −80°C until analysis by MS.

For tEGFR, purified receptor was supplied in a buffer containing 0.03% dodecylmaltoside and 0.5 mM DTT at a concentration of ~10 μM (as determined by the Leahy lab by SDS-PAGE and comparison to BSA standards). To gain insight into the upper limit of coverage in the extracellular coverage of tEGFR, a 6 μl aliquot quenched with 20 μl QB (0.1 % FA pH 2.8 6.2 M Guanidine 500 mM TCEP) for ten minutes at room temperature then diluted with FB (0.8% FA 10% glycerol) to make final 180 mM TCEP 2 M Gndn-HCl and final volume for each reaction 60 μl.

Protein digestion, peptide fragmentation and MS

H/DX samples were individually thawed at 0°C for 2 min, then injected (50 µl) and pumped through an immobilized pepsin (Sigma) column at al flow rate of 100 µl/min for 3 minutes. Pepsin (Sigma) was immobilized by amine coupling to POROS 20 AL support (Applied Biosystems) and packed into column housings of 2 mm × 2 cm (64 µl) (Upchurch). Protease-generated fragments were collected onto a C18 HPLC trap column (800 µm × 2 mm, Dionex). Peptides were eluted into and through an analytical C18 HPLC column (0.3 × 75 mm, Agilent) by a nonlinear 10–50% buffer B gradient over 15 minutes at 9 µl/min (Buffer A: 0.1% formic acid; Buffer B: 0.1% formic acid, 99.9% acetonitrile) followed by a 10 minute wash cycle that increased the % Buffer B to 95%. Each injection was also followed immediately by an additional 30 minute wash protocol that cycled from 10% Buffer B to 80% Buffer B four times using a flow rate of 9 µl/min. The effluent was electrosprayed into the mass spectrometer (LTQ Orbitrap XL, Thermo Fisher Scientific). The SEQUEST (Bioworks v3.3.1) software program (Thermo Fisher Scientific) was used to identify the likely sequence of the parent peptides using non-deuterated samples via tandem MS using a database that contained EGFR as well as several other systems being studied by the Englander group. A comprehensive peptide pool for wild-type sEGFR was developed by four iterative rounds of

MS/MS using rejection and inclusion lists of previously identified peptides, and the same peptide pool was used for all subsequent analyses of wild-type sEGFR with and without EGF.

The mass spectrometer used was a ThermoScientific LTQ Orbitrap XL (Thermo Fisher Scientific, Waltham, MA, USA) operated at a 60,000 resolution (~1 s/scan). Data were collected in profile mode with an AGC target of 106. The mass calibration was checked daily and recalibrated when necessary (G2 ppm rms deviation over 9 masses/100 scans in the calibration mix). Source parameters were: spray voltage 3.5 kV; capillary voltage 40 V; tube lens 170 V; capillary temperature 150 °C. MS/MS CID fragment ions were usually detected in the LTQ stage of the instrument in centroid mode at normal scan rate with an AGC target value of 104. CID fragmentation was at 35% energy for 30 ms at Q of 0.25.

H/DX data analysis

MATLAB-based MS data analysis tool—ExMS—was used for data processing, using a retention time window of 3 minutes (Kan, 2011). Briefly, the ExMS program searches raw MS data, identifies individual isotopic peaks/envelopes from a list of MS/MS peptides obtained from SEQUEST search and calculates centroid values of these envelopes. The program is used to first identify the isotopic envelope centroid and chromatographic elution time of each parental non-deuterated peptide, and then this information is subsequently used to identify deuterated peptides.

CHAPTER THREE FIGURES

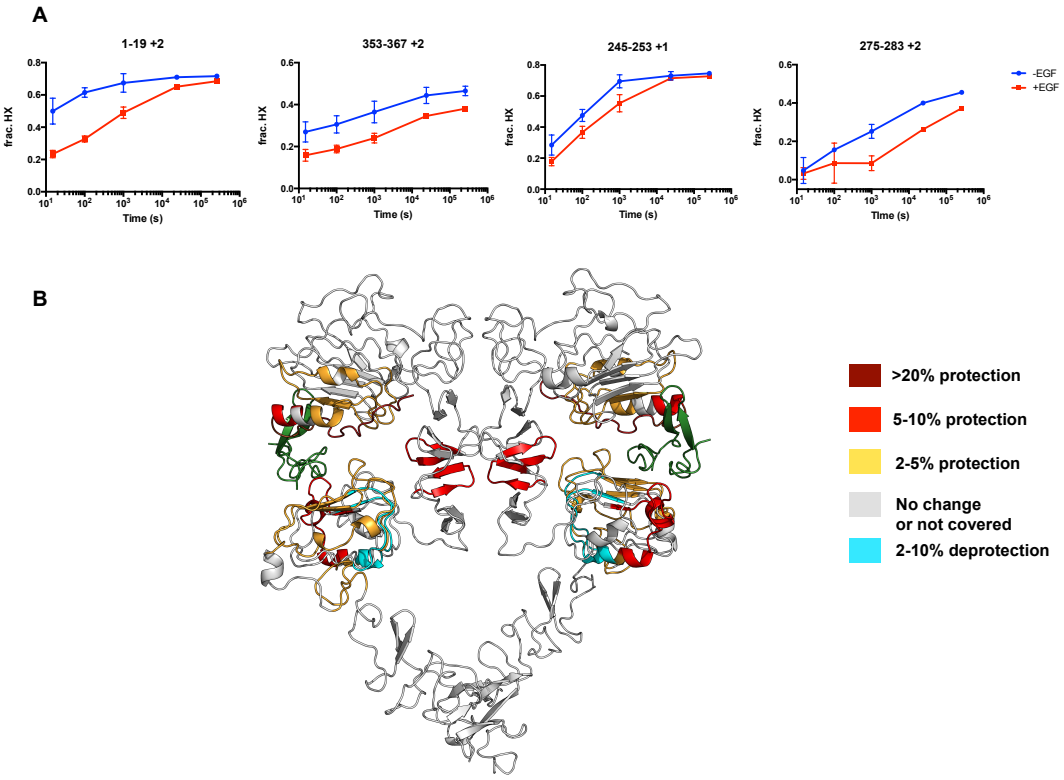


Figure 3.1 Preliminary HDX/MS analysis of the extracellular region of EGFR. (A) Data for representative peptides in the ligand binding and dimerization interfaces in the EGFR extracellular region. Amide hydrogen/deuterium exchange rates calculated by normalizing the centroid shift of the isotopic envelope compared to the monoisotopic mass by the maximum number of exchangeable amides for a given peptide. Plots are plotted over a timecourse (in seconds), on a log scale, with standard deviations of independent replicates of each timepoint indicated by the error bars. The first two peptides on the left are in domains I and III, respectively, and are peptides that are involved in direct interactions with EGF. The two peptides on the right are in the dimerization interface. 245-253 is the β hairpin that forms the direct 'tether' interaction, while 275-283 forms the 'buttressing' interaction. **(B)** % change

from the 100 sec timepoint (unless indicated) plotted as binned on the structure of the EGF (shown in green) bound sEGFR dimer (PDB ID 3NJP). The only exception to this (to highlight the small deprotection at longer timescales) is the % deprotection (shown in cyan) of the aa. 370-380 loop. The % deprotection here is calculated from the longest timepoint (t ~10^5 s).

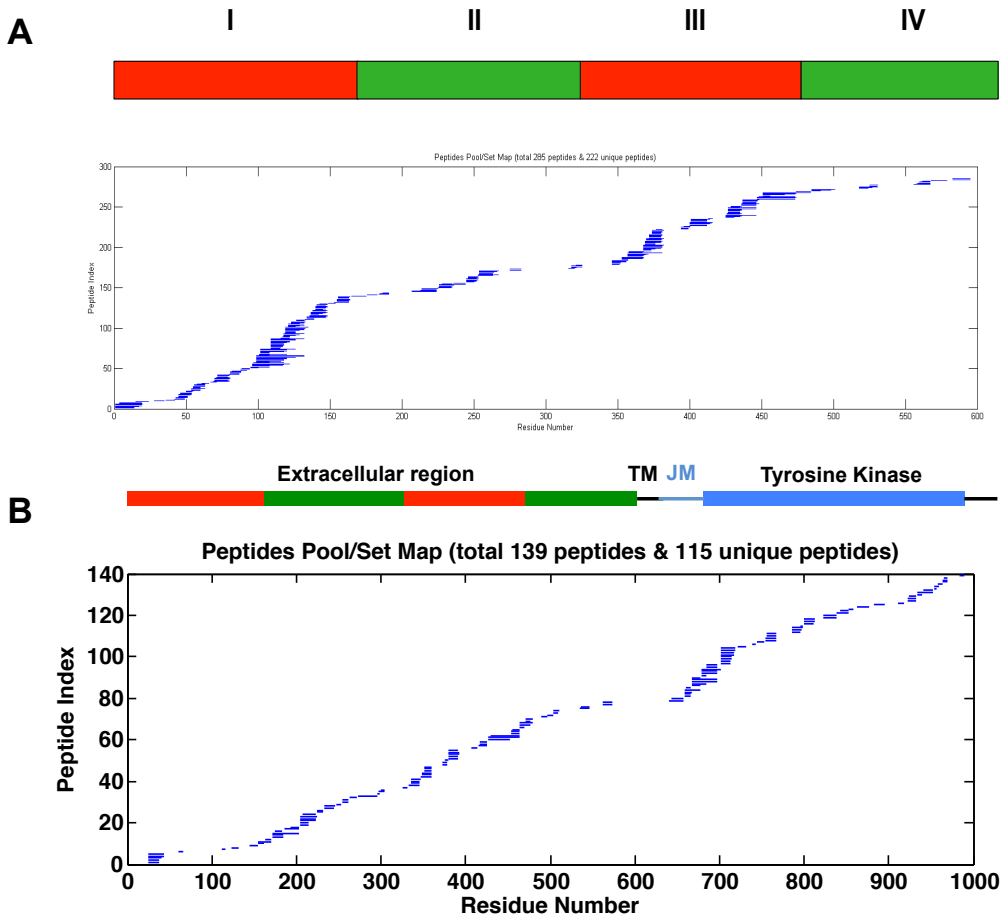


Figure 3.2 Preliminary MS/MS coverage of EGFR extracellular region and tEGFR (full length EGFR aa. 1-998) in dodecylmaltoside micelles. (A) MS/MS coverage of the extracellular region of EGFR compared to the domain architector of sEGFR. Coverage is limited in the cysteine rich domains shown in green. Each blue bar represented one peptide mapped onto the amino acid residue number for the sEGFR construct, for which a domain

architecture is shown in the top part of the panel. **(B)** MS/MS coverage of intact tEGFR in dodecylmaltoside micelles.

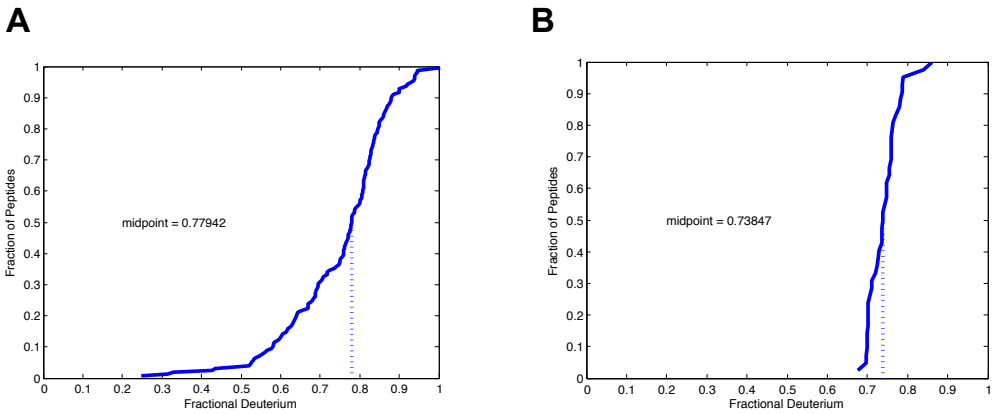


Figure 3.3 Distribution of fractional deuteration of 'all-D' sample of sEGFR compared to cytochrome C. (A) Distribution of fractional deuteration of peptides in the optimal 'all D' sample of sEGFR under the conditions described in the Methods. The fractional deuteration is a result from both experimental deuteration as well as back exchange and is compared to a positive control (cytochrome C). The half-maximal fractional deuteration across all peptides is indicated by the dashed line. There is a long tail to the left of this line that are peptides in the domain I and III solenoid structures that are difficult to unfold under these conditions. **(B)** A positive control (cytochrome C) for evaluation of the distribution with the same quench protocol used for analysis of sEGFR.

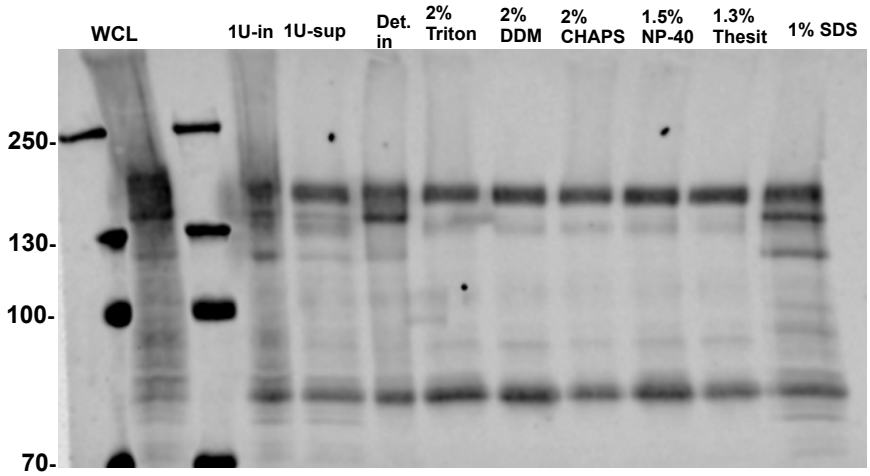


Figure 3.4 Assessment of FL-EGFR detergent extraction efficiency. FL-EGFR was purified as described in the Methods section from Sf9 cells. Whole cell lysate (WCL), prepared by direct lysis of Sf9 cells in reducing Laemmli buffer, was compared quantitatively to assess protein loss at early stages of the purification and detergent extraction efficiency. 1U-in represents the input to the 1st ultracentrifugation step, and 1U-sup is the supernatant. As can be seen, a large fraction of the protein is lost at this step, suggesting this step may need more optimization. Det-in. is the input to the detergent extraction step, and compared quantitatively to six different detergents indicated. As can be seen, the detergent extraction efficiency is quite high, and does not vary significantly across detergents.

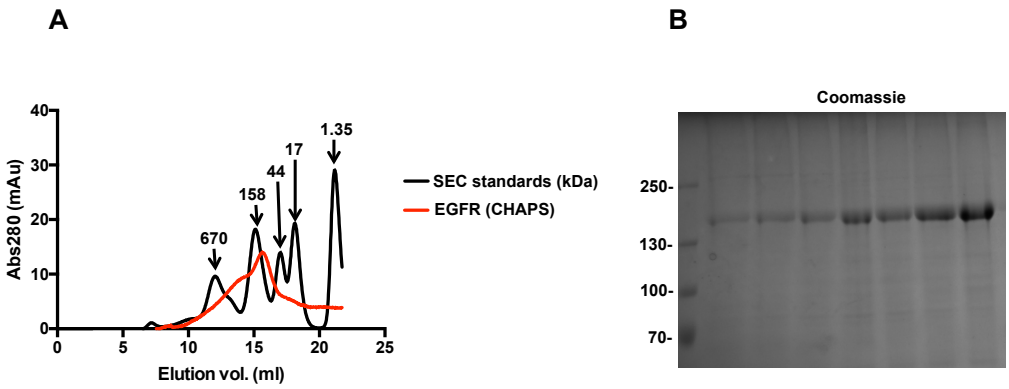


Figure 3.5 Progress towards purification of full length EGFR. The CHAPS-solubilized full-length EGFR, purified as described in the Methods, was examined by size-exclusion chromatography using a Superose 6 column (GE Healthcare). As can be seen, the ~170 kDa protein elutes mostly as a monomer by SEC, although the stokes radius may be slightly lower than would be predicted for a monomer, suggesting it may be compact in the context of a CHAPS micelle. There is also an aggregated 'shoulder' at lower elution volumes. (B) Fractions containing the peaks shown in the aborbance trace shown in panel A were concentrated 10-fold by TCA precipitation and visualized by SDS-PAGE and Coomassie staining. As shown, the full length EGFR is >80% pure.

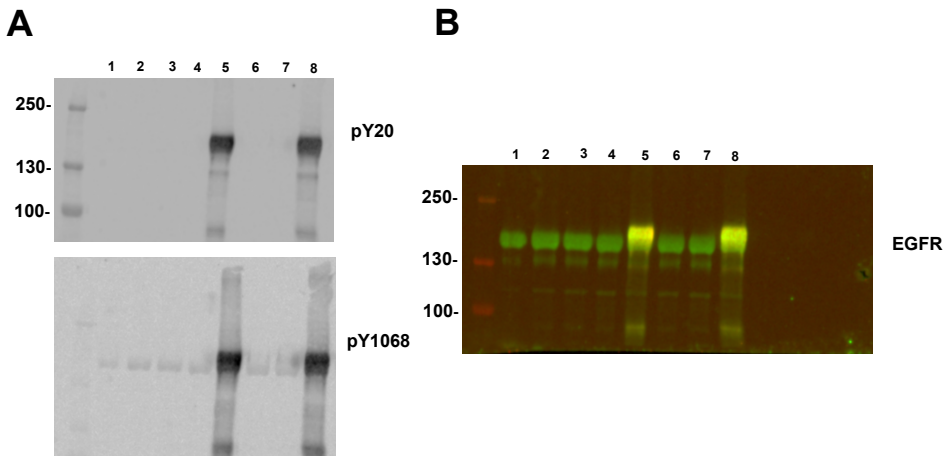


Figure 3.6 Preliminary autophosphorylation analysis of full-length EGFR in CHAPS. Autophosphorylation reactions on Talon and ion exchange purified FL-EGFR were done as stated in the Methods, and analyzed by SDS-PAGE and immunoblotting using antibodies specified in Methods. Lane order: (order of 10 mM $MgCl_2$, 1 mM ATP, 40 μM EGF). 1. - - - 2. + - - 3. - - + 4. - + - 5. + + - 5. + - + 7. - + + 8. + + +

REFERENCES

Adak, S., Yang, K.S., Macdonald-Obermann, J., and Pike, L.J. (2011). The membrane proximal intracellular domain of the Epidermal Growth Factor receptor underlies negative cooperativity in ligand binding. J Biol Chem *286*, 45146-45155.

Alvarado, D., Klein, D.E., Lemmon, M.A. (2009). ErbB2/HER2/Neu resembles an autoinhibited invertebrate EGF receptor. Nature *461*, 287-291.

Bessman, N. J., Bagchi, A., Ferguson, K.M., & Lemmon, M.A. (2014). Complex relationship between ligand binding and dimerization in the epidermal growth factor receptor. Cell Reports *9*, 1306-1317.

Bobst, C. E., & Kaltashov, I.A. (2014). Enhancing the quality of H/D exchange measurements with mass spectrometry detection in disulfide-rich proteins using electron capture dissocation. Anal Chem *86*, 5225-5231.

Burke, J. E., Karbarz, M.J., Deems, R.A., Li, S., Woods, V.L. Jr., & Dennis, E.A. (2008). Interaction of Group IA Phospholipase A2 with Metal Ions and Phospholipid Vesicles Probed with Deuterium Exchange Mass Spectrometry. Biochemistry *47*, 6451-6459.

Chen, J., Rempel, D.L., & Gross, M.L. (2010). Temperature jump and fast photochemical oxidation probe submillisecond protein folding. J Am Chem Soc *132*, 15502-15504.

Chung, K. Y., Rasmussen, S.G., Liu, T., Li, S., DeVree, B.T., Chae, P.S., Calinski, D., Kobilka, B.K., Woods, V.L. Jr., & Sunahara, R.K. (2011). Conformational changes in the G protein Gs induced by the beta2 adrenergic receptor. Nature *477*, 611-615.

Cohen, S., Carpenter, G., & King, L. Jr. (1980). Epidermal growth factor-receptor-protein kinase interactions. Co-purification of receptor and epidermal growth factor-enhanced phosphorylation activity. J Biol Chem *255*, 4834-4842.

Coskun, U., Crzybek, M., Drechsel, D., & Simons, K. (2011). Regulation of human EGF receptor by lipids. Proc Natl Acad Sci USA *108*, 9044-9048.

Dawson, J. P., Berger, M.B., Lin, C.-C., Schlessinger, J., Lemmon, M.A., and Ferguson, K.M. (2005). Epidermal growth factor receptor dimerization and activation require ligand-induced conformational changes in the dimer interface. Mol Cell Biol *25*, 7734-7742.

Dawson, J. P., Bu, Z., and Lemmon, M.A. (2007). Ligand-induced structural transitions in ErbB receptor extracellular domains. Structure *15*, 942-954.

Duc, N. M., Du, Y., Thorsen, T.S., Lee, S.Y., Zhang, C., Kato, H., Kobilka, B.K., & Chung, K.Y. (2015). Effective application of bicelles for conformational analysis of G protein-coupled receptors by hydrogen/deuterium exchange mass spectrometry. J Am Soc Mass Spectrom *26*, 808-817.

Ferguson, K. M., Berger, M., Mendrola, J., Cho, H.S., Leahy, D.J., and Lemmon, M.A. (2003). EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. Molecular Cell *11*, 507-517.

Ferguson, K. M., Darling, P. J., Mohan, M. J., Macatee, T. L., Lemmon, M. A. (2000). Extracellular domains drive homo- but not hetero-dimerization of erbB receptors. EMBO J *19*, 4632-4643.

Freed, D. M., Alvarado, D.A., & Lemmon, M.A. (2015). Ligand regulation of a constitutively dimeric EGF receptor. Nature Communications *6*, 7380.

Hu, W., Kan, Z.-Y., Mayne, L., & Englander, S.W. (2016). Cytochrome c folds through foldon-dependent native-like intermediates in an ordered pathway. Proc Natl Acad Sci USA *in press.*

Kan, Z.-Y., Mayne, L., Chetty, P.S., & Englander, S.W. (2011). ExMS: data analysis for HX-MS experiments. J Am Soc Mass Spectrom *22*, 1906-1915.

Kan, Z.-Y., Walters, B.T., Mayne, L., & Englander, S.W. (2013). Protein hydrogen exchange at residue resolution by proteolytic fragmentation mass spectrometry analysis. Proc Natl Acad Sci USA *110*, 16438-16443.

Lu, C., Mi, L.Z., Grey, M.J., Zhu, J., Graeuf, E., Yokoyama, S., and T.A. Springer (2010). Structural evidence for loose linkage between ligand binding and kinase activation in the epidermal growth factor receptor. Mol Cell Biol, 5432-5443.

Lu, C., Mi, L.Z., Shurpf, T., Walz, T., and Springer, T.A. (2012). Mechanisms for kinase-mediated dimerization of the epidermal growth factor receptor. J Biol Chem *287*, 38244-38253.

Macdonald-Obermann, J. L., & Pike, L.J. (2009). The intracellular juxtamembrane domain of the epidermal growth factor (EGF) receptor is responsible for the allosteric regulation of EGF binding. J Biol Chem *284*, 13570-13576.

Mi, L. Z., Lu, C., Li, Z., Nishida, N., Walz, T., and T.A. Springer (2011). Simultaneous visualization of the extracellular and cytoplasmic domains of the epidermal growth factor receptor. Nat Struct Mol Biol *18*, 984-989.

Qiu, C., Tarrant, M.K., Boronina, T., Longo, P.A., Kavran, J.M., Cole, R.N., Cole, P.A., & Leahy, D.J. (2009). In vitro enzymatic characterization of near full length EGFR in activated and inhibited states. Biochemistry *48*, 6624-6632.

Red Brewer, M., Choi, S.H., Alvarado, D., Moravcevic, K., Pozzi, A., Lemmon, M.A., and Carpenter, G. (2009). The juxtamembrane region of the EGF receptor functions as an activation domain. Molecular Cell *34*, 641-651.

Ronan, T., Macdonald-Obermann, J.L., Huelsmann, L., Bessman, N.J., Naegle, K.M., & Pike, L.J. (2016). Different Epidermal Growth Factor Receptor (EGFR) Agonists Produce Unique Signatures for the Recruitment of Downstream Signaling Proteins. J Biol Chem *291*, 5528-5540.

Shukla, A. K., Westfield, G.H., Xiao, K., Reis, R.I., Luang, L.Y., Tripathi-Shukla, P., Qian, J., Li, S., Blanc, A., Oleskie, A.N., Dosey, A.M., Su, M., Ligang, C.R., Gu, L.L., Shan, J.M., Chen, X., Hanna, R., Choi, M., Yao, X.J., Klink, B.U., Kahsai, A.W., Sighu, S.S., Koide, S., Penczek, P.A., Kossiakoff, A.A., Woods, V.L., Jr., Kobilka, B.K., Skiniotis, G., Lefkowitz, R.J. (2014). Visualization of arrestin recruitment by a G-protein coupled receptor. Nature *512*, 218-222.

Skinner, J. J., Lim, W.K., Bédard, S., Black, B.E., Englander, S.W. (2012). Protein dynamics viewed by hydrogen exchange. Protein Science *21*, 996-1005.

Vivanco, I., Robins, H.I., Rohle, D., Campos, C., Grommes, C., Nghiemphu, P.L., Kubek, S., Oldrini, B., Chheda, M.G., Yannuzzi, N., Tao, H., Zhu, S., Iwanami, A., Kuga, D., Dang, J., Pedraza, A., Brennan, C.W., Heguy, A., Liau, L.M., Liberman, F., Yung, W.K., Gilbert, M.R., Reardon, D.A., Drappatz, J., Wen, P.Y., Lamborn, K.R., Chang, S.M., Prados, M.D., Fine, H.A., Horvath, S., Wu, N., Lassman, A.B., DeAngelis, L.M., Yong, W.H., Kuhn, J.G., Mischel, P.S., Mehta, M.P., Cloughesy, T.F., & Mellinghoff, I.K. (2012). Differential sensitivy of glioma- verses lung cancer-specific EGFR mutations to EGFR kinase inhibitors. Cancer Discovery *2*, 458-471.

Walters, B. T., Ricciuti, A., Mayne, L., & Englander, S.W. (2012). Minimizing back exchange in the hydrogen exchange-mass spectrometry experiment. J Am Soc Mass Spectrom *23*, 2132-2139.

Wang, Z., Longo, P.A., Tarrant, M.K., Kim, K., Head, S., Leahy, D.J., and Cole, P.A. (2011). Mechanistic insights into the activation of oncogenic forms of EGF receptor. Nat Struct Mol Biol *18*, 1388-1393.

Wang, Z., Raines, L.L., Hooy, R.M., Roberson, H., Leahy, D.J., & Cole, P.A. (2013). Tyrosine phosphorylation of mig6 reduces its inhibition of the epidermal growth factor receptor. ACS Chem Biol *8*, 2372-2376.

Wood, E. R., Truesdale, A.T., McDonald, O.B., Yuan, D., Hassell, A., Dickerson, S.H., Ellis, B., Pennisi, C., Home, E., Lackey, K., Alligood, K.J., Rusnak, D.W., Gilmer, T.M., & Shewchuk, L. (2004). A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells. Cancer Research *64*, 6652-6659.

Yarden, Y., Schlessinger, J. (1987). Epidermal growth factor induces rapid, reversible aggregation of the purified epidermal growth factor receptor. Biochemistry *26*, 1443-1451.

Yu, X., Sharma, K.D., Takahashi, T., Iwamoto, R., and Mekada, E. (2002). Ligand independent dimer formation of epidermal growth factor receptor (EGFR) is a step separable from ligand-induced EGFR signaling. Mol Cell Biol *13*, 2547-2557.

Zhang, X., Gureasko, J., Shen, K., Cole, P.A., and Kuriyan, J. (2006). An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. Cell *125*, 1137-1149.

CHAPTER FOUR

Molecular basis of Necitumumab/Portrazza™ inhibition of EGFR variants associated with acquired cetuximab resistance.

SUMMARY

The epidermal growth factor receptor (EGFR) family is implicated in a number of cancers and members of the family are the targets of monoclonal antibody therapies. Cetuximab/Erbitux™ is an EGFR antibody that is approved for head and neck and colorectal cancers. Unfortunately, resistance to this targeted therapy limits its therapeutic benefit. Our collaborators at Eli Lilly show that Necitumumab/Portrazza™, recently approved by the U.S. Food and Drug Administration for squamous non small cell lung cancer, inhibits EGFR with cetuximab and panitumumab resistance mutations. To understand how necitumumab can bind to resistance mutations despite sharing a nearly identical epitope, I determined a 2.8 Å X-ray crystal structure of the necitumumab Fab fragment in complex with isolated EGFR domain III bearing the S492R (S468R in mature numbering) cetuximab resistance mutation. Comparison of the paratopes of cetuximab and necitumumab to those of other therapeutic antibodies reveals a structural class of antibodies that contain buried cavities between their VH and VL domains. My structural analysis reveals that plasticity and hydration of the necitumumab paratope may make it a useful therapeutic agent for these cases of cetuximab resistance and for future combination therapies. We also identify paratope shape as a potential feature of antibodies that may be utilized in synergistic modes in combinatorial therapeutics.

SIGNIFICANCE

Resistance to anti-epidermal growth factor receptor monoclonal antibody therapies, including cetuximab/Erbitux™ is often driven by epitope mutations. Necitumumab/Portrazza™ binds and inhibits EGFR mutants that cause cetuximab resistance despite sharing a common

epitope, through structural plasticity of its paratope, which I observe in an X-ray crystal structure of EGFR domain III S492R (S468R in mature numbering) in complex with the Fab fragment from necitumumab. Necitumumab therefore may be a useful agent in future combination therapies.

INTRODUCTION

The Epidermal Growth Factor (EGF) receptor or ErbB/HER family of receptor tyrosine kinases (RTKs) have been implicated in and drive many human cancers. It is well established that antibodies to the extracellular region of the EGF receptor (EGFR) that prevent ligand binding and conformational changes required for receptor activation can inhibit tumor growth *in vivo*. Therapeutic antibodies in current clinical use that target the ErbB family include cetuximab/Erbitux™, panitumumab/Vectibix™ – both of which target EGFR/HER1 itself—and trastuzumab/Herceptin™, which targets ErbB2/HER2. In addition, there are several antibodies that target ErbB3/HER3 in clinical trials, including KTN3379 (Kolltan Pharmaceuticals, New Haven, CT) (Lee, 2015).

Cetuximab/Erbitux™ is in active clinical use to treat colorectal, and head and neck squamous cell cancers. As with all targeted cancer therapies, acquired resistance to cetuximab limits the duration of effective treatment. In some cases resistance to cetuximab has been attributed to mutations in intracellular signaling molecules, such as *KRAS* and *BRAF*. In 2012, a mutation in extracellular region of EGFR itself was identified in metastatic KRAS wild-type colorectal cancer (mCRC) tumors that had acquired resistance to cetuximab treatment (Montagut et al., 2012). These alterations gives rise to a single amino acid substitution in the cetuximab epitope (S492R in pro-EGFR or S468R, using the numbering system that starts from the mature protein). In that study it was reported that another EGFR

targeted monoclonal antibody drug, panitumumab, is an effective inhibitor of the S468R mutant and therefore would be the therapy of choice for cetuximab resistant patients.

Recent exome sequencing analysis of *KRAS* wild-type colorectal cancer tumors has systematically identified genomic alterations in response to EGFR cetuximab therapy associated with resistance to cetuximab (Arena et al., 2015; Bertotti, 2015; Medico, 2015; Morelli, 2013; Siravegna, 2015) which include genetic alterations in *EGFR, ERBB2, PDGFR*, and amplification of *FGFR1, MET,* and *ERBB2* (Bardelli, 2013; Yonesaka, 2011). Additional EGFR extracellular region mutations - K443T, G441E, G441R, and I467M (in mature amino acid numbering) – were also identified following treatment with cetuximab, and with panitumumab, in patient samples and in cell lines (Table 1) (Bertotti, 2015; Briag, 2015; Siravegna, 2015). Whereas panitumumab was shown to bind and inhibit K443T and S468R, it has greatly diminished binding for I467M and G441E/R. G441R is also associated with resistance to both cetuximab and panitumumab in gastrointestinal cancer (Briag, 2015). The G441E EGFR mutant, when ectopically expressed in cetuximab-sensitive cell line NCI-H508 cells, reduces cetuximab EGFR inhibition and increases activation of downstream signals in a manner that is dependent on expression of the downstream adaptor IRS2, consistent with previous observations that these mutations are sufficient in inducing resistance to cetuximab therapy (Bertotti, 2015).

Overall, EGFR mutations within the shared epitope of cetuximab and panitumumab represent a major mechanism of clinical resistance to cetuximab; 16 % of mCRC patients in the cetuximab arm of the ASPECCT trial developed the S468R mutation, where only 1 % of patients did so in the panitumumab arm (Newhall, 2014; Price, 2015), suggesting that that this resistance mutation occurs more selectively after EGFR inhibition by cetuximab.

Necitumumab/Portrazza™ (IMC-11F8), an anti-EGFR IgG1 antibody therapy recently approved in combination with gemcitabine and cisplatin for squamous non small cell lung

carcinoma, shares a very similar epitope to that of cetuximab (IMC-225) and panitumumab, though the paratopes are quite different (Li, 2008; Li et al., 2005). Like cetuximab, its epitope overlaps with the EGF binding site, and therefore binding of necitumumab interferes with binding of EGF to EGFR (Li, 2008).

The X-ray crystal structures of the necitumumab Fab fragment bound to isolated EGFR domain III (sEGFR-d3) and the full length EGFR extracellular region (sEGFR) were previously determined and show that necitumumab has completely different binding interactions from cetuximab despite sharing a nearly identical epitope (Li, 2008). Previous analysis also identified a similar shape complementarity parameter in the necitumumab paratopic interface compared to that of cetuximab, despite having a buried hydrophobic cavity between its light and heavy chains, where it interacts with EGFR domain III (Li, 2008) .

In this chapter I report that necitumumab retains high affinity for four described EGFR epitope mutations that cause cetuximab resistance. Whereas these EGFR mutations greatly diminish the binding affinities of cetuximab, their effects are drastically smaller on binding of necitumumab. Our collaborators at Ely Lilly show that necitumumab effectively inhibits EGFR bearing the most commonly observed of these mutations – the S468R mutation – and that necitumumab causes dose-dependent EGFR-S468R internalization and degradation.

I also report the X-ray crystal structure to 2.8 Å resolution of the Fab fragment from necitumumab (Fab11F8) in complex with isolated domain III from EGFR (sEGFRd3) with the serine at position 468 substituted with arginine (sEGFRd3-S468R). I was able to resolve clearly the side chain of the arginine at position 468 and to visualize how it is accommodated in this cavity in the interface between EGFR domain III and the Fab, and additional interactions that stabilize it, including a water-mediated hydrogen bond and several cation-pi interactions.

My structural analysis reveals that the packing between the light and heavy chains and consequent flatness of cetuximab paratope leave it susceptible to alterations by genetic mutation. By contrast necitumumab contains a hydrophobic and hydrated cavity capable of exhibiting structural plasticity. I also observe minor conformational changes within the necitumumab paratope, which may contribute to such plasticity. Moreover, I suggest that the distinct modes of binding of cetuximab and necitumumab, despite sharing a common epitope, can potentially be exploited in the context of combinatorial therapies.

RESULTS

Necitumumab binds with high affinity and inhibits EGFR bearing cetuximab resistance mutations.

To ask if necitumumab is able to bind with high affinity to EGFR mutants that cause cetuximab resistance, we collaborated with Eli Lilly and Company (New York, NY; Indianapolis, IN) to analyze antibody binding to HEK293T expressing EGFR bearing cetuximab resistance mutation S468R by immunostaining with Alexa-647 labeled antibodies. We find that necitumumab and panitumumab can retain binding and inhibition to cells that overexpress EGFR harboring the S468R (EGFR-S468R) mutation whereas cetuximab does not (Eli Lilly and Company, data, not shown). Treatment of cells with necitumumab leads to inhibition and dose dependent downregulation and degradation of both wild type and EGFR-S468R, whereas the EGFR-S468R is not effectively inhibited by cetuximab, with an associated decrease in receptor downregulation (Eli Lilly and Company, data, not shown).

The arginine side chain at position 468 (S468R) in cetuximab resistant patients would clash with the cetuximab paratope. The serine at position 468, however, projects towards a buried hydrophobic cavity (Figure 4.2) in the necitumumab paratope, which led us to

hypothesize that necitumumab would retain EGFR binding affinity. The presence of this cavity makes no significant difference between measures of shape complementarity between cetuximab and necitumumab (Table 4.4). Increasing the packing density in this cavity by replacing this serine with an isoleucine (S468I) enhances necitumumab affinity while not altering cetuximab affinity (Li, 2008).

The presence of this hydrophobic cavity (Figure 4.2) in this central location in the EGFR/necitumumab interface prompted us to hypothesize that it may bind to other EGFR mutants that cause cetuximab resistance in this region of the cetuximab epitope on EGFR. We find that necitumumab retains high affinity to four cetuximab resistance EGFR mutations. Two of these mutations directly project into this cavity (S468R/I467M), one is adjacent (S440L), and one is peripheral to the paratopic interface (K443T), suggesting additional structural considerations underlie the ability of necitumumab to bind, and therefore inhibit, these EGFR mutants (Figure 1).

Binding of sEGFR harboring cetuximab resistance mutations to the Fab fragments of cetuximab and necitumumab.

To evaluate the extent of diminished binding to cetuximab of the soluble extracellular region of EGFR (sEGFR) harboring mutations that have been linked to cetuximab resistance we used surface plasmon resonance (SPR/Biacore), exactly as we have done in the past (Ferguson et al., 2003; Li et al., 2005).

We generated sEGFR variants with the patient derived cetuximab resistance mutations S468R and K443T and with the cell line derived mutations S440L and I467M. Proteins with alterations at G441 could not be generated. As shown in Figure 4.3B and Table 4.2, sEGFR-S468R shows very weak binding (K_D value >> 4 μM) to immobilized FabC225, and alteration at the adjacent amino acids (sEGFR-I467M) reduces binding by more than 50-fold. Substitution of leucine at position S440, which is located close to S468 on the adjacent β-strand of domain III, also dramatically reduced binding to FabC225 (K_D value >> 4 μM), and substitution of K443 with threonine reduced binding more than 100-fold (Figure 4.3A). Of note, we had previously observed that substitution of K443 with alanine, did not affect measurably cetuximab binding (Li, 2008). These same sEGFR variants bind to immobilized EGF with very similar affinities to those observed for wild type sEGFR, indicating that the loss of binding affinity for cetuximab is not due to a general destabilization of domain III of the protein but to the local effect of these particular amino acid substitutions.

We next quantified the binding energetics of EGFR bearing cetuximab resistance mutations by surface plasmon resonance of recombinantly expressed and purified mutants of the EGFR extracellular region. We engineered all of these mutations in the recombinant baculovirus expressed extracellular region of EGFR (sEGFR). sEGFR binds to the Fab fragment of necitumumab (Fab11F8) and cetuximab (FabC225) with high nanomolar affinity (Table 4.2). We find that binding of sEGFR-S468R to immobilized Fab11F8 is just 2-fold weaker than binding of wild type sEGFR, in stark contrast to the dramatic loss of binding of this protein to FabC225 (Figure 4.3C). A variant of sEGFR harboring the other patient derived mutation (sEGFR-K443T) shows similar high affinity binding to immobilized Fab11F8, whereas the two cell line derived variants, sEGFR-I467M and sEGFR-S440L, show slightly weaker binding with K_D values of 27 and 53 nM respectively. For these cetuximab resistance sEGFR variants, binding of necitumumab is predicted to be 10 to ~ 500 fold stronger than binding of cetuximab (Table 4.2).

X-ray crystal structure of sEGFRd3-S468R/Fab11F8.

To gain molecular insight into the ability of necitumumab to bind with high affinity and inhibit these sEGFR variants we determined the 2.8 Å resolution X-ray crystal structure of the necitumumab Fab fragment bound to domain III of EGFR harboring the S468R cetuximab resistance mutation (sEGFRd3-S468R/Fab11F8). X ray diffraction data collection and refinement statistics are shown in Table 4.3.

Clear electron density could be seen for the arginine side chain at amino acid position 468 in initial molecular replacement solutions. Following refinement, clear density can be seen for the side chain in each of the four molecules in the asymmetric unit (Figure 4.4), as well as an ordered coordinated water molecule in three of the four complexes. Electron density for both the arginine side chain and coordinated water molecule can be seen when both are omitted from the model (Figure 4.4) and in simulated annealing composite omit maps (data not shown). The structure reveals that necitumumab is able to accommodate R468 in a buried hydrophobic cavity between the light and heavy chains that is hydrated.

In addition to being sterically accommodated, R468 makes several important interactions that contribute to stabilizing the introduction of a basic residue in a hydrophobic environment. The guanidinium moiety in the arginine side chain engages in hydrogen bonds with the main chain carbonyl of G100A in CDR H3 and a water mediated hydrogen bond with the side chain of S35 in CDR H1 (using Chothia numbering for the VH amino acids). Hydration of this cavity is essential for stabilizing R468 in a relatively hydrophobic environment. In addition to a water mediated interaction, R468 engages in cation-pi interactions with F100C, Y33, and Y50. The Y50 sidechain rotates ~15 degrees towards the R468 to contribute to this interaction (Figure 4.1B). Views of electrostatic potential of the

necitumumab paratope alone (Figure 4.2) reveals that F100C contributes to a region of electronegative potential within the hydrophobic cavity, consistent with it being able to contribute to a cation-pi interaction.

The structures of EGFR domain III wildtype and domain III S468R are very similar -- domain III overlays with an pairwise root mean square displacement (RMSD) of 0.458 Å, while overlaying the entire structure of the complex including Fab11f8 overlay with an RMSD of 0.569 Å, indicating that the EGFR S468R mutation does not dramatically alter the conformation of either EGFR or the Fab fragment from necitumumab. The mean buried interface area in the domain III S468R complex is very similar – 896 $Å^2$, as compared to 875 $Å^2$ in the wild-type structure, consistent with the view that the mutation fills a small cavity that is otherwise unoccupied.

One notable difference in the mutated structure is that the binding of the necitumumab Fab to this mutated EGFR domain III is less rigid than in the wild-type structure, as evidenced by a structural 'wobble.' Upon overlaying domain III (aa 311-480) of EGFR in each of the four complexes in the asymmetric unit, we noticed a 'wobble' in the necitumumab light chain conformation. The conformation of the VL domain are more different in the four complexes in the asymmetric unit in the domain III S468R structure than they are the eight complexes in the asymmetric unit in the domain III wild-type structure. Analysis of these different conformations reveals a screw motion of the antibody complex, with a small translation away from the site of mutation and rotation around the two-fold symmetry axis within the Fab (Figure 4.5).

Comparison of the necitumumab and cetuximab paratopes to those of other therapeutic antibodies.

As mentioned previously the R468 side chain in sEGFRd3-S468R is accommodated in a hydrophobic cavity between the VL and VH chains in this Fab. Overlaying the

sEGFRd3S468R from our structure with the domain III from the structure of sEGFR bound to FabC225 immediately shows that there is insufficient space to accommodate the arginine due to a steric clash with the side chain of Y100A (in Chothia numbering) in CDR H3 (Fig. 4.2) of cetuximab. A surface view comparison of the cetuximab paratope reveals that it is relatively flat, where the necitumumab paratope has a buried cavity between its VH and VL domains. Comparison of the necitumumab and cetuximab paratopes further reveals that this difference in paratope shape is due to CDR H3 conformation. In cetuximab, the conformation of CDR H3 sterically blocks the ability of this antibody to accommodate amino acid substitutions in EGFR located at the apex of domain III that the antibody paratopes target (Figure 4.1). This sterically 'locked' conformation is stabilized by bivalent polar interactions between D100 in the VH domain and N91 and Y50 in the VL domain (Figure 4.2).

The presence of the hydrophobic cavity in necitumumab makes no significant difference in quantitative measures of shape complementarity (Lawrence, 1993) (Table 4.4). Identifying this cavity in necitumumab capable accommodation of alterations in its target prompted us to compare the paratopes of cetuximab and necitumumab to those of other fourteen other therapeutic antibodies for which complex structures have been determined (Table 4.4). Analysis of sixteen different antibody paratopes revealed a structural class of antibodies that contain buried cavities similar to the one found in necitumumab. My analysis did not reveal a significant correlation between the shape complementarity (Sc) and the presence of a buried cavity between the light and heavy chains, nor between the antibody origin and shape complementarity. However, there is a greater ratio of mean interface area to shape complementarity in antibodies that contain buried cavities in their paratopes, presumably because the antibody must bury more surface area in order to form a cavity. Other examples of this trend include MORO9825, KTN3379, and RG7116, all anti-ErbB3 antibodies that have an Sc statistic of 0.62-0.65. Comparison of the ratio of buried surface area to Sc of KTN3379 to the ratios of MORO9825, and RG7116 reveals that antibodies with

higher ratios tend to have larger buried cavities akin to the one found in necitumumab. While there is no physical meaning to this ratio, in combination with contact analysis (MacCallum, 1996), it may serve as a potentially useful benchmark for evaluating new antibody paratopes in terms of the presence of similar buried cavities akin to that of necitumumab and other therapeutic antibodies. Further studies are needed to evaluate whether other such cavities in this structural class of antibodies would be less susceptible to resistance through epitope mutations such as the ones studied here.

DISCUSSION

Resistance to targeted cancer therapeutics remains a formidable challenge that has increased the necessity for the development of combination therapies, and sequential therapeutic rechallenge approaches. Epitope mutations, a major mechanism of cetuximab resistance in KRAS wild-type colorectal cancer, induce resistance by decreasing the affinity of cetuximab for EGFR. High affinity cetuximab binding to EGFR is necessary for its ability to block ligand binding and conformational changes and therefore EGFR activation. In the case of cetuximab, high affinity EGFR binding is driven through relatively high shape complementarity and the presence of a high abundance of bulky aromatic residues in the CDR regions. Presence of bulky aromatic amino acid sidechains in complementarity determining regions (CDR's) and high shape complementarity is often thought necessary to promote high affinity binding (Chatellier, 1996; Ohno, 1985), and there have been several efforts to increase the prevalence of such bulky residues in therapeutic antibody CDR's (Fisher, 2010).

I find here that necitumumab is an effective inhibitor of EGFR mutants that cause resistance to cetuximab, despite sharing a common epitope. My structural analysis reveals

that hydration of a hydrophobic cavity in the necitumumab paratope is critical for its effectiveness in accommodating genetic changes to its epitope. The presence of this cavity is driven by the conformation of CDR H3, which is known to be hypervariable and 'kinked' (Weitzner, 2015; Xu, 2015) (Figure 4.2). CDR H3 may be a marker for presence of such a buried cavity more generally. Necitumumab also exhibits structural plasticity in binding less rigidly to a genetically altered epitope (Figure 4.5). We find through biochemical analysis that it is less sensitive to genetic alterations in its epitope on the EGFR than cetuximab, and may be an effective therapeutic agent in the context of cetuximab resistance through epitope mutations (Figure 4.3).

I observe in an X-ray crystal structure of the S468R mutation in domain III of EGFR bound to the Fab fragment of necitumumab that hydration of this cavity, as well as several cation-pi interactions in the necitumumab paratope may help to stabilize the S468R mutation in this hydrophobic pocket. It is well established that solvation and desolvation of antibody/antigen interfaces can play important roles in both enthalpic and entropic contributions to antibody/antigen interaction energetics (Acchione, 2009; Bhat, 1994). Buried water molecules at antibody/antigen interfaces, in order to contribute favorably to the thermodynamics of antibody/antigen binding, must mitigate the inherent entropic penalty associated with limiting the degrees of freedom of the water molecules through enthalpic contributions through hydrogen bonding networks.

In the context of the necitumumab interaction with EGFR-S468R, the energetic penalty associated with burial of a positively charged residue in a region of low dielectric constant would likely be higher than the 2-fold decrease in binding we observed for necitumumab binding to sEGFR S468R, supporting the thermodynamic importance of the buried water molecule in providing a hydrogen-bond acceptor for the buried guanidinium of the arginine sidechain, which is an important enthalpic contribution that stabilizes the buried positive charge in this cavity. Such a trapped water molecule may still, however, lead to an entropic

penalty in the thermodynamics of antibody binding to EGFR, and therefore necitumumab binding to the S468R mutation could potentially be engineered to be even greater if a direct hydrogen bond with R468 were engineered. The presence of this hydrophobic cavity facilitates accommodation of genetic mutations that can cause resistance to antibodies that target the same epitope in the EGFR, such as cetuximab and panitumumab.

While cetuximab and necitumumab are no different in quantitative measures of shape complementarity, the method of Colman and Lawrence to quantitate protein/protein interface shape complementarity relies on a subset selection of vector normals to surfaces calculates in GRASP (Lawrence, 1993). This calculation, while excellent for evaluating general shape complementarity and has been adopted as the most common tool to do so, may not be the appropriate calculation for evaluating the presence of a cavity at a protein/protein interface due to its exclusion of longer vectors that would penetrate into a buried cavity. However, Thornton and colleagues have developed a computational approach to measure relative 'concavity' of antibody paratopes by contact analysis, which may be a more useful measure for distinguishing between antibodies that are capable of exhibiting a similar mode of structural plasticity and those that cannot (MacCallum, 1996).

Although it has been previously reported that panitumumab is also an effective inhibitor of the S468R EGFR mutant, there are some differences between cetuximab and panitumumab in terms of their mechanism of action. Cetuximab (IgG1) and panitumumab (IgG2) both compete with binding of EGF and other EGFR ligands to EGFR domain III, therefore inhibiting receptor activation and signaling. It is also clear through the case of trastuzumab/Herceptin, an ErbB2-directed antibody, that ErbB-targeted antibodies have an important role in mediating antibody-dependent cellular toxicity through engaging Fc receptors (Arteaga, 2012; Baselga, 2001; Molina, 2001; Spector, 2009). It is also clear that cetuximab and panitumumab both also have an ADCC component that is important to their therapeutic effects, and the precise consequences of ADCC are different for IgG1 vs. IgG2

antibodies (Schneider-Merck, 2010). In addition, some of the resistance mutations also cause a decrease in binding affinity for panitumumab, like I467M (Arena et al., 2015). There is no X ray crystal structure of panitumumab bound to EGF receptor, but it is possible that this loss of affinity is due to a similar steric clash seen in cetuximab.

There are several developing approaches to overcome mechanisms of resistance to anti-EGFR antibodies involving mutations in the antibody epitopes, most commonly antibody mixtures of distinct epitopes, of which there are several in clinical trials. One such mixture called Sym-004, is a cocktail of two anti-EGFR monoclonal antibodies with non-overlapping epitopes on EGFR domain III (Pedersen, 2010) This antibody mixture can overcome resistance to cetuximab in a cell line derived from non-small cell lung cancer NCI-H226 with induced cetuximab resistance, leading to EGFR inhibition and degradation (Iida, 2013). Data from a phase I trial of Sym-004 suggest potent activity in mCRC patients. However, this study did also find that the presence of the S468R mutation in one patient diminished activity of Sym-004, which could suggest one of antibodies may also be susceptible to this mutation like cetuximab. Another similar combination therapy, called Pan-HER, is a combination of six synergistic antibodies that simultaneously target epitopes on EGFR, HER2, and HER3 (Jacobsen, 2015). This antibody combination is effective at inducing growth inhibition of several cancer cell lines and xenograft models, including ones with acquired resistance to cetuximab, suggesting that simultaneous targeting of several ErbB/HER family members may be an effective approach to overcome therapeutic resistance (Jacobsen, 2015).

Both approaches illustrate the value and effectiveness of evolving combinatorial therapeutics, and evaluation of the appropriate synergistic combinations of antibodies for such strategies remains an important avenue of investigation. They also illustrate that most commonly exploited functional property of therapeutic antibodies in the context of combination therapies is the epitope. Our analysis shows for the first time that concavity and shape of an antibody paratope is another such property that can be exploited as a functional

synergism. We find through comparison of sixteen antibody complex structures that the ratio of buried surface area to shape complementarity may be a useful approximate predictor of such a cavity to complement existing contact analysis algorithms (MacCallum, 1996). These tools may serve as useful descriptive measures for antibody paratope shape for identification of synergistic pairs of antibody paratopes for combination therapies in the context of resistance to just one therapeutic.

METHODS

Protein production.

The secreted full length extracellular region of EGFR (sEGFR; amino acids [aa] 1-618 of mature EGFR), and of isolated domain III of EGFR (sEGFRd3; aa 1-4 followed by 311-514) were produced in a baculovirus system from Sf9 cells as described previously (Ferguson et al., 2000; Li et al., 2005). Cetuximab resistance variants were created using standard PCR methods and produced exactly as for wild type proteins. The sEGFR and variants were purified by Ni-NTA chromatography, followed by size-exclusion chromatography (SEC) as previously described (Ferguson et al., 2000). For sEGFRd3, an additional cation exchange purification step was included prior to SEC. Cetuximab Fab fragment (FabC225) and necitumumab IgG were provided by Imclone Systems Corporation (New York, NY). The necitumumab Fab (Fab11F8) was generated by papain digestion using a Pierce Fab Preparation Kit. Following removal of the Fc on a protein A column, the Fab was further purified on a Superose 12 size exclusion chromatography column (GE Healthcare) equilibrated in 25 mM HEPES, 100 mM NaCl (pH 7.5).

Surface Plasmon Resonance binding studies

Surface Plasmon Resonance (SPR) experiments were carried out on a Biacore 3000 instrument at 25°C in buffer containing 25 mM HEPES, pH 8.0, 150 mM NaCl, 3 mM EDTA, and 0.005% nonidet-p20 (HBS-EP8) exactly as previously described (Li et al., 2005). FabC225 and Fab11F8, at 50 μg/ml in 10 mM sodium acetate, pH 5.5, were separately amine coupled to activated CM5 sensor surfaces by 5 minutes exposure at 10 μl/minute. EGF (Intergen Inc.) was immobilized at 200 μg/ml in 10 mM sodium acetate, pH 4.0. A series of samples of the relevant sEGFR or sEGFRd3 at different concentrations in HBS-EP 8 buffer were injected over these surfaces and the equilibrium SPR response relative to a blank control surface (blocked with ethanolamine) was measured. Injections with contact times of 25 minutes at 10 μl/min for Fab surfaces and 15 minutes at 5 μl/min for EGF surface were used to ensure complete equilibration at all protein concentrations. Regeneration between samples was used to restore baseline SPR response: 1 minute of 10 mM glycine-HCl, 1 M NaCl, pH 2.5 for Fab surfaces and 1 minute of 10 mM sodium acetate, 1 M NaCl, pH 4.5 for EGF surface. For wild type sEGFR the equilibrium SPR response for each sample was plotted against protein concentration, and the curve fit to a one-site Langmuir binding model using Prism 6 (GraphPad Software, Inc.). The B_{max} value from these fits for wild type sEGFR binding to each ligand were then used to normalize the equilibrium SPR responses for binding of sEGFR variants. These fraction maximal SPR binding values were plotted against concentration and fit to simple one-site Langmuir binding model with the B_{max} fixed to 1.0. Experiments were conducted at least three times and the mean K_D value determined from the fits to each independent, normalized binding curve (Table 4.2).

Crystallization, data collection and structure determination

For crystallization, samples containing a 1:2 molar ratio of sEGFRd3-S468R and Fab11F8 preequilibrated at room temperature for one hour were applied to a Superose 6 size exclusion column (GE Healthcare) that had been equilibrated in 25 mM HEPES, 200 mM

NaCl, pH 7.5. Fractions containing the sEGFRd3-S468R/Fab11F8 complex were pooled, concentrated to 5-8 mg/ml and crystallized using the hanging drop vapor diffusion method at 20° C. Initial needle-like crystals of sEGFRd3-S468R/Fab11F8 were obtained from drops comprising 0.5 µl complex in 25 mM HEPES, 100 mM NaCl, pH 7.5 plus 0.5 µl of reservoir solution containing 50 mM sodium acetate, 250 mM ammonium sulfate, 15-20% PEG3350, pH 5.0. Crystals appeared within one month. Diffraction to 3 Å resolution was obtained from clusters of such needles but single crystals suitable for data collection could not be obtained. Large hexagonal plate crystals were grown using streak seeding with the needle clusters into drops containing 0.5 µl complex at 5-7 mg/ml in 25 mM HEPES 200 mM NaCl, pH 7.5 and 0.5 µl reservoir solution containing 50 mM sodium acetate, 200 mM potassium citrate, 15-20% PEG3350, pH 6-6.3. Crystals appeared within 2-3 weeks, and continued to grow for several months. These crystals were flash frozen in liquid nitrogen following brief exposure to a cryoprotectant of reservoir solution supplemented with 12 % ethylene glycol. Crystals were of space group $P2_1$. Data were collected at the GM/CA APS beamline 23-ID-D using a Pilatus 6M detector. The data were processed in HKL2000 (Otwinowski and Minor, 1997). Data collection statistics are summarized in Table 3.

The structure of sEGFRd3-S468R/Fab11F8 was solved by the method of molecular replacement (MR) using the program PHASER (McCoy, 2007). Domain III (aa 311-502) and Fab11F8 (aa 1-210 from light chain and 1-221 heavy chain) from PDB ID 3B2U were used as independent search models to identify the four sEGFRd3-S468R/Fab11F8 complexes in the asymmetric unit. Clear electron density for the arginine side chain at position 468 could be seen in initial molecular replacement maps. Initial stages of refinement used rigid body groups to refine the positions of the Fab variable domains. Further manual refinement proceeded with iterative rounds of manual model rebuilding in COOT (Emsley, 2010) and refinement in PHENIX (Adams et al., 2010). Later stages of refinement included TLS refinement using each chain as a TLS group. To check for possible model bias, simulated

annealing composite omit maps were calculated in PHENIX and CNS (Brunger et al., 1998) using torsion angle dynamics at 2500 K.

Structure Analysis.

Structures were overlaid and RMSD values were calculated using the programs PYMOL (Delano Scientific, Palo Alto, CA), or COOT (Emsley, 2010), and reflect main chain atoms only. Shape complementarity values were calculated using the Sc module in CCP4 (Lawrence, 1993), excluding saccharide, ligand, and solvent molecules. Buried interface surface area was calculated using the PISA server. Most of the X ray diffraction and structural analyses were completed using software compiled on SBGrid (Morin, 2013).

CHAPTER FOUR TABLES

Table 4.1.

Mutation	Alternate numbering	Cetuximab treated patients	Panitumumab treated patients	Cetuximab exposed cell line
S440L	S464L			X
G441R	G465R	X	X	X
G441E	G465E	X	X	
K443T	K467T	X		
I467M	I491M			X
S468R	S492R	X		X

Table 4.1. Acquired antibody resistance mutations in the cetuximab/panitumumab epitope of EGFR (Arena et al., 2015; Bertotti et al., 2015; Montagut et al., 2012; Siravegna, 2015).

Table 4.2.

Immobilized Ligand	K_D Values for binding to sEGFR (nM)				
	WT	**S440L**	**I467M**	**K443T**	**S468R**
FabC225	5.7 ± 0.5	>> 4000	296 ± 50	670 ± 103	>> 4000
Fab11F8	6.4 ± 0.9	53.7± 7.3	27.4 ± 1.5	13.8 ± 0.4	12.4 ± 1.4
EGF	147 ± 16	93.7± 10	175.9 ± 10.9	256 ± 25	107 ± 12

Table 4.2. Mean K_D values for binding of sEGFR wild type and cetuximab resistance mutations to immobilized Fab fragments of cetuximab (FabC225) and necitumumab (Fab11F8) and to EGF.

Table 4.3.

Table 4.3. X-ray diffraction data collection and refinement statistics

	sEGFRd3-S468R/Fab11F8
Data collection	
Space group	$P2_1$
Cell dimensions	
a, *b*, *c* (Å)	151.9, 80.8, 172.9
α, β, γ (°)	90, 91.5, 90
Resolution (Å)	2.8 (2.9-2.79)
R_{merge}	0.18 (0.72)
$I / \sigma I$	8.4 (1.2)
Completeness (%)	96.4 (94.3)
Redundancy	4.1 (2.8)
Refinement	
Resolution (Å) (d_{min})	50-2.8 (2.83-2.79)
No. unique reflections	99,558
R_{work} / R_{free}	0.236/0.256 (0.352/0.367)
No. atoms	
Protein	19,545
Ligand/ion	28
Waters	142
B-factors	
Protein	68.06
Ligand/ion	
Water	
R.m.s. deviations	
Bond lengths (Å)	0.002
Bond angles (°)	0.677
Ramachandran favored (%)	95.6
Ramachandran outliers (%)	0.4
Rotamer outliers (%)	0.4

Table 4.4. Structural features of selected therapeutic antibody/antigen interfaces.

Therapeutic antibody (target antigen)	**PDB ID**	**Source**	**A.S.U.**	**Sc statistic (Lawrence, 1993)**	**Interface area (Å^2)**	**Mean ratio of interface area to Sc (Å^2)**	**Cavity**
Necitumu-mab/IMC-11F8 (EGFR domain III wild-type)	3B2U	Humanized from chimeric library	8	0.68± 0.03	878.1± 29.3	1292.5	Y
Necitumu-mab/IMC-11F8 (EGFR domain III S468R)	T.B.D.	Humanized from chimeric library	4	0.69± 0.02	895.9± 54.8	1300.3	Y
Cetuximab/ IMC-225 (EGFR)	1YY9	Chimeric	1	0.703	875	1244.7	N
Sifalimumab (IFN-α2A)	4YPG	Undisclosed	2	0.676± 0.004	1260.3± 0.03	1864.3	Y
Fab 3379 (ErbB3)	5CUS	Undisclosed	4	0.654± 0.03	835.9± 12.7	1279.1	I
MORO9825 (ErbB3)	4P59	Phage display selected from human combinatorial antibody library (HuCAL GOLD)	1	0.626	1140.3	1821.6	Y
(aP2)	5C0N	Cloned and purified from rabbits	2	0.612± 0.05	456.8± 12.8	746.9	N

		immunized with recombinant human and mouse aP2					
(aP2)	5D8J	Cloned and purified from rabbits immunized with recombinant human and mouse aP2	1	0.722	721.7	999.6	N
(Jag1)	5Bo1	Phage display selected with recombinant human JAG1 and JAG2 from mouse and human from an undisclosed library	2	0.75± 0.02	992± 15	1323.7	Y
Onartuzumab (c-MET)	4K3J	Hybridomas from mice immunized with recombinant Met IgG and subsequently humanized and affinity matured	1	0.772	831.1	1076.6	Y
Fab 39.29 (Influenza hemaglutin A)	4KVN	Cloned and selected from mice immunized with influenza strain combination	1	0.761	1058.2	1390.5	I

RG7116 (ErbB3)	4LEO	Cloned from mice immunized with recombinant ErbB3, and subsequently humanized and glycoengineered	1	0.64	927	1448.4	Y
Infliximab (TNFα)	4G3Y	Undisclosed	1	0.629	941.1	1496.2	I
Canakinumab (IL-1β)	4G6J	human IgGκ monoclonal antibody of undisclosed origin	1	0.726	950.6	1309.3	I
Basiliximab (IL-2Rα)	3IU3	Chimeric mouse-human antibody	3	0.561± 0.01	1023.4±67.3	1825.4	Y
(Notch1 NRR)	3L95	Phage display selected from human antibody library	2	0.715± 0.02	871.2±20.6	1219.3	N
Daclixumab (IL-2Rα)	3NFP	Undisclosed	2	0.586± 0.001	916.6±78.2	1564.2	I

CHAPTER FOUR FIGURES

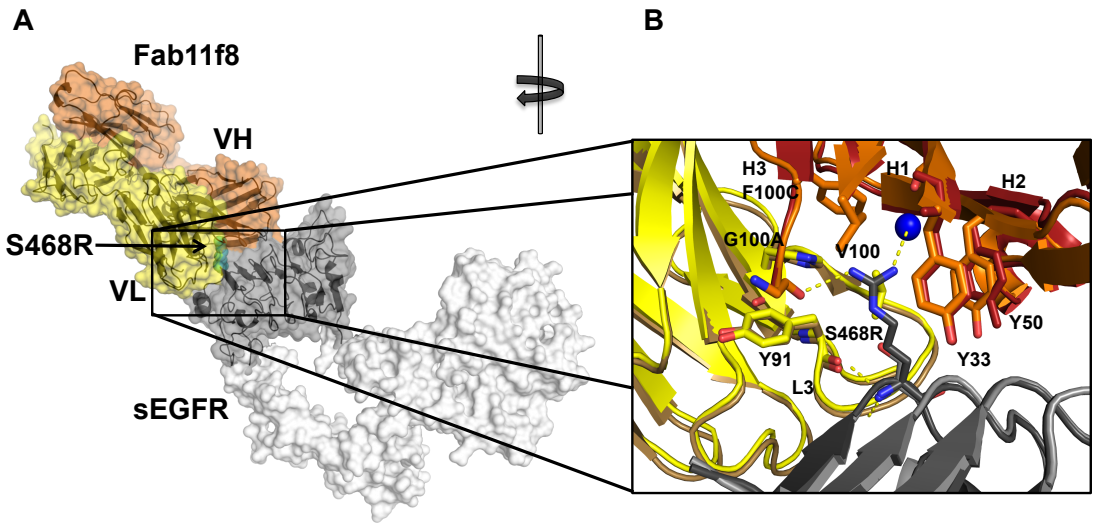


Figure 4.1 Structural basis for Necitumumab inhibition of EGFR-S468R.

Necitumumab accommodates the S468R cetuximab resistance mutation in a buried hydrophobic cavity between its VH and VL domains that is hydrated in the structure of sEGFRd3-S468R bound to the Fab fragment of necitumumab (Fab11f8). This buried water molecule, coordinated by S35 (in Chothia numbering), in addition to a hydrogen bond with the backbone carbonyl from G100A in the VL domain, and several cation-pi interactions due to the proximity of R468 to F100C, Y33 and Y50 in the VH domain, serve to stabilize the basic guanidinium moiety of R468 in this region of low dielectric constant.

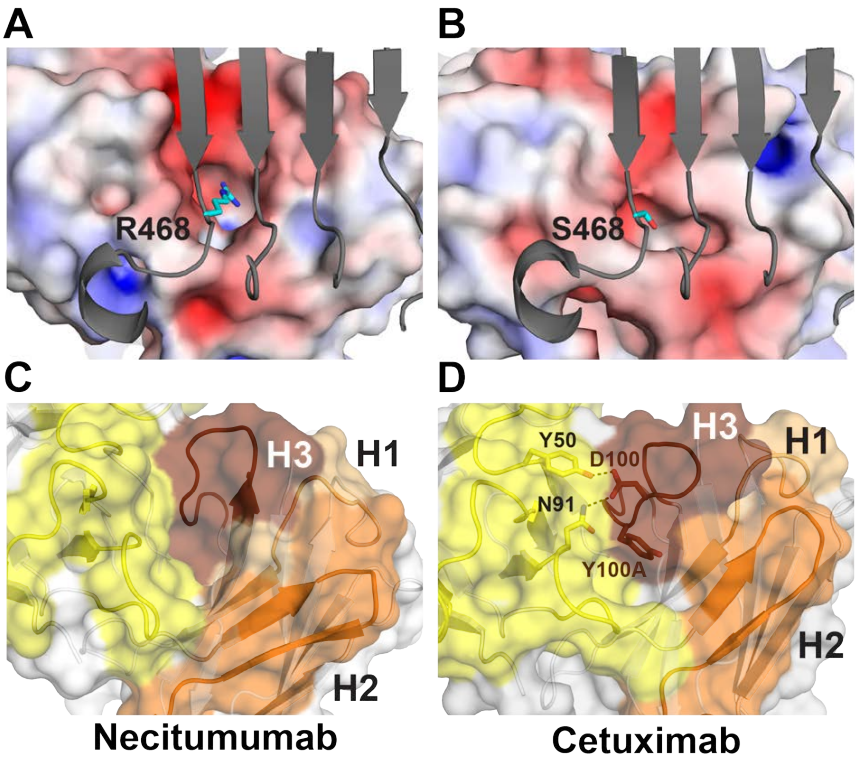


Figure 4.2. EGFR-S468R interacts with a hydrophobic cavity in necitumumab with high electronegativity. (A) & (B) Electrostatic potential calculated in APBS (Baker, 2001) and mapped onto the surfaces of the respective antibody paratope surfaces show that cetuximab and necitumumab have similar electrostatic potential properties of their paratopes. In both antibodies, there is a region of relatively high electronegative potential close to the region of the EGFR-S468R mutation. **(C) & (D)** Comparison of the paratopes of necitumumab and cetuximab reveals that the presence of the hydrophobic cavity in necitumumab is a consequence of CDR H3 conformation. In cetuximab, the CDR H3 conformation may be stabilized by several polar interactions, and a salt bridge between D100 and N91, as a consequence poising the sidechain of Y100A to clash with the site of mutations associated with acquired cetuximab resistance.

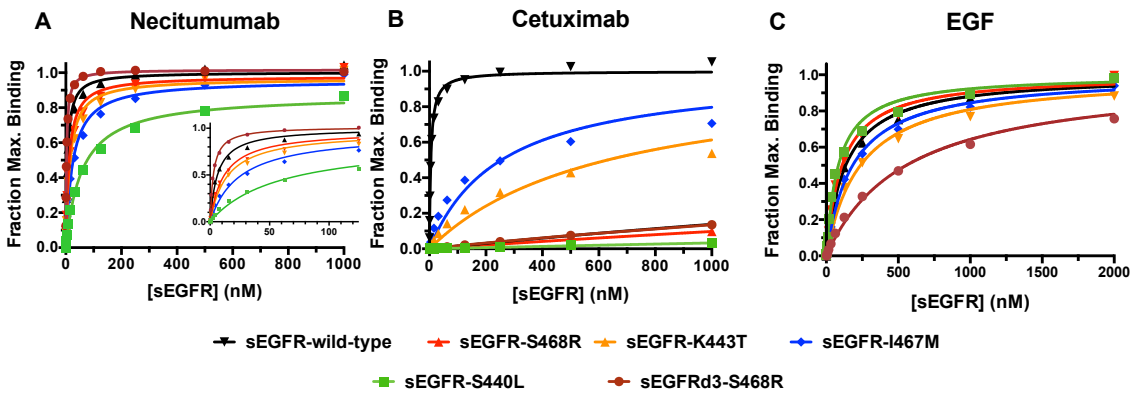


Figure 4.3. Necitumumab binds with high affinity to EGFR mutations that cause cetuximab resistance. Samples of the indicated concentrations of sEGFR and variants were passed over Biacore surfaces to which (A) Fab11f8 (Necitumumab), (B) FabC225 (Cetuximab) and (C) EGF had been immobilized. Fabs and EGF were immobilized on Biacore CM5 chips using amine coupling. Binding assay and data processing were conducted as described in the experimental procedures. Representative, normalized binding curves are show and mean K_D values are reported in Table 2.

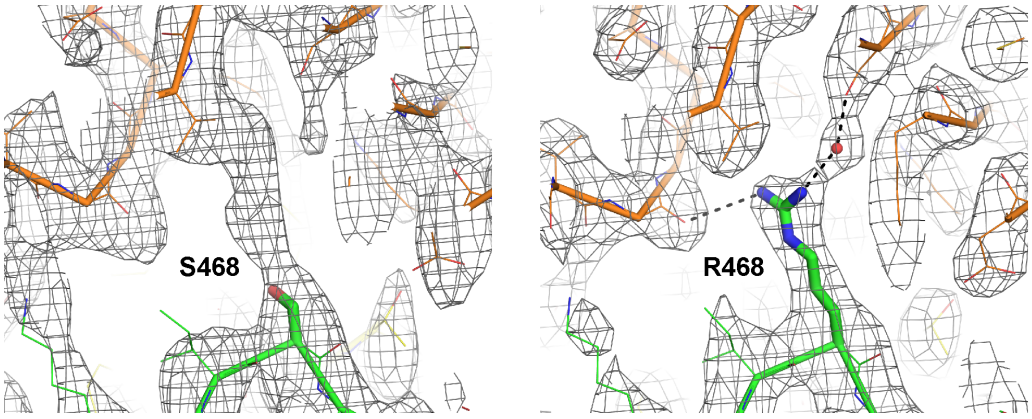


Figure 4.4. Electron density around one example of R468 in sEGFRd3-S468R/Fab11F8 structure.

(A) MR model is shown with domain III in green, VH in orange and VL in yellow. Main chain ribbon is shown and side chains are shown as lines. S468 is shown as in stick representation. A $2mF_o$-DF_c map from this PHASER model is shown contoured at 1.0 σ. **(B)** Current model is shown, displayed as in A with the corresponding $2mF_o$-DF_c map.

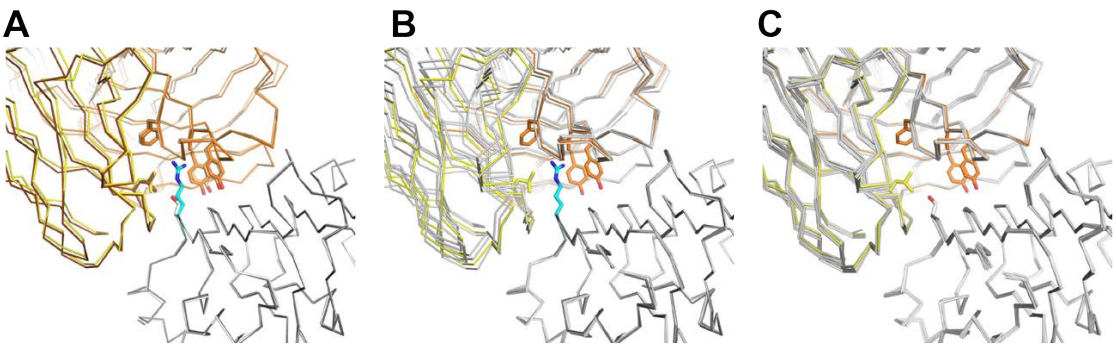


Figure 4.5. Necitumumab exhibits a structural 'wobble' in binding to EGFR-S468R.

(A) Comparison of the necitumumab complexes of wildtype sEGFRd3 and sEGFRd3-S468R, using only the coordinates of domain III to overlay the structures. Domain III is shown in grey with the side chain at position 468 shown in stick representation; S468 in yellow and R468 in cyan. The Fab is in yellow (VL) and pale orange (VH) when bound to wild type sEGFRd3 and in brown (VL) and orange (VH) when bound to sEGFRd3-S468R. Minor differences can be discerned in the relative positions of the VH and VL chains and in the positions of several side chains. **(B)** A similar domain III superposition of the 4 examples of the sEGFRd3-S468R/Fab11F8 complexes in the asymmetric unit show that there are significant differences in the positions of the Fab with respect to domain III in each copy of the complex. For the complex show in A VH in orange and VL in yellow. The other 3 complexes are show in grey. **(C)** A similar overlay of all 8 complexes of sEGFRd3/Fab11F8 from PDB ID 3B2U shows that the 8 structures are much more similar.

REFERENCES

Acchione, M., Lipschultz, C.A., DeSantis, M.E., Shanmuganathan, A., Li, M., Wlodawer, A., Tarasov, S., & Smith-Gill, S.J. (2009). Light chain somatic mutations change thermodynamics of binding and water coordination in the HyHEL-10 family of antibodies. Mol Immunol *47*, 457-464.

Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W*., et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr *D66*, 213-221.

Arena, S., Bellosillo, B., Siravegna, G., Martinez, A., Canadas, I., Lazzari, L., Ferruz, N., Russo, M., Misale, S., Gonzalez, I*., et al.* (2015). Emergence of Multiple EGFR Extracellular Mutations during Cetuximab Treatment in Colorectal Cancer. Clinical cancer research : an official journal of the American Association for Cancer Research *21*, 2157-2166.

Arteaga, C. L., Sliwkowski, M.X., Osborne, C.K., Perez, E.A., Puglisi, F., & Gianni, L. (2012). Treatment of HER2-positive breast cancer: current status and future perspectives. Nature Reviews Clinical Oncology *9*, 16-32.

Baker, N. A., Sept, D., Joseph, S., Holst, M.J., & McCammon, J.A. (2001). Electrostatics of nanosystems: Applications to microtubules and the ribosome. Proc Natl Acad Sci USA *98*, 10037-10041.

Bardelli, A., Corso, S., Bertotti, A., Hobor, S., Valtorta, E., Siravegna, G., et al. (2013). Amplication of the MET receptor drives resistance to anti-EGFR therapies in colorectal cancer. Cancer Discovery *3*, 658-673.

Baselga, K., Albanell, J., Molina, M.A., & Arribas, J. (2001). Mechanism of action of trastuzumab and scientific update. Semin Oncol *28*, 4-11.

Bertotti, A., Papp, E., Jones, S., Adleff, V., Anagnostou, V., Lupo, B., Sausen, M., Phallen, J., Hruban, C. A., Tokheim, C*., et al.* (2015). The genomic landscape of response to EGFR blockade in colorectal cancer. Nature *526*, 263-267.

Bertotti, A., Papp, E., Jones, S., Adleff, V., Anagnostou, V., Lupo, B., Sausen, M., Phallen, J., Hruban, C.A., Tokheim, C., Niknafs, N., Nesselbush, M., Lytle, K., Sassi, F., Cottino, F., Migliardi, G., Zanella, E.R., Ribero, D., Russolillo, N., Mellano, A., Muratore, A., Paraluppi, G., Salizzoni, M., Marsoni, S., Kragh, M., Lantto, J., Cassingena, A., Li, Q.K., Karchin, R., Scharpf, R., Sartore-Bianchi, A., Siena, S., Diaz, L.A. Jr., Trusolino, L., Velculescu, V.E. (2015). The genomic landscape of response to EGFR blockade in colorectal cancer. Nature *526*, 263-267.

Bhat, T., Bentley, G.A., Boulot, G., Greene, M.I., Tello, D., Dall'Acqua, W., Souchon, H., Schwarz, F.P., Mariuzza, R.A. & Roljak, R.J. (1994). Bound wate rmolecules and conformational stabilization help mediate an antigen-antibody association. Proc Natl Acad Sci USA *91*, 1089-1093.

Briag, F., et al. (2015). Epidermal growth factor receptor mutation mediates cross resistance to panitumumab and cetuximab in gastrointestinal cancer. Oncotarget *6*, 12035-12047, doi: 12010.18632/oncotarget.13574.

Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S.*, et al.* (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta crystallographica Section D, Biological crystallography *54 (Pt 5)*, 905-921.

Chatellier, J., Van Regenmortal, M.H., Vernet, T., & Altschuh, D. (1996). Functional mapping of conserved residues located at the VL and VH domain interface of an Fab. J Mol Biol *264*, 1-6.

Emsley, P., B. Lohkamp, W.G. Scott, and K. Cowtan (2010). Features and Development of Coot. Acta Crystallographica D *66*, 486-501.

Ferguson, K. M., Berger, M. B., Mendrola, J. M., Cho, H. S., Leahy, D. J., and Lemmon, M. A. (2003). EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. Mol Cell *11*, 507-517.

Ferguson, K. M., Darling, P. J., Mohan, M. J., Macatee, T. L., and Lemmon, M. A. (2000). Extracellular domains drive homo- but not hetero-dimerization of erbB receptors. EMBO J *19*, 4632-4643.

Fisher, R. D., Ultsch, M., Lingel, A., Schaefer, G., Shao, L., Birtalan, S., Sidhu, S.S., & Eigenbrot, C. (2010). Structure of the comlex between HER2 and an antibody paratope formed by side chains from tryptophan and serine. J Mol Biol *402*, 217-229.

Iida, M., Brand, T.M., Starr, M.M., Li, C., Huppert, E.J., Luthar, N., Pedersen, M.W., Horak, I.D., Kragh, M., Wheeler, D.L. (2013). Sym004, a novel EGFR antibody mixture, can overcome acquired resistance to cetuximab. Neoplasia *15*, 1196-1206.

Jacobsen, H. J., Poulsen, T.T., Dahlman, A., Kjaer, I., Koefoed, K., Sen, J.W., Weilguny, D., Bjerregaard, B., Andersen, C.R., Horak, I.D., Pedersen, M.W., Kragh, M., and Lantto, J. (2015). Pan-HER, an Antibody Mixture Simultaneously Targeting EGFR, HER2, and HER3 Effectively Overcomes Tumor Heterogeneity and Plasticity. Clinical Cancer Research *21*, 4110-4122, doi: 4110.1158/1078-0432.CCR-4114-3312.

Lawrence, M. C., Colman, P.M. (1993). Shape complementarity at protein/protein interfaces. J Mol Biol *234*, 946-950.

Lee, S., Greenlee, E.B., Amick, J.R., Ligon, G.F., Lillquist, J.S., Natoli, E.J. Jr., Hadari, Y., Alvarado, D., Schlessinger, J. (2015). Inhibition of ErbB3 by a monoclonal antibody that locks the extracellular domain in an inactive configuration. Proc Natl Acad Sci USA *112*, 13225-13230.

Li, S., Kussie, P., Ferguson, K.M. (2008). Structural basis for EGF receptor inhibition by the therapeutic antibody IMC-11F8. Structure *16*, 216-227, doi: 210.1016/j.str.2007.1011.1009.

Li, S., Schmitz, K. R., Jeffrey, P. D., Wiltzius, J. J., Kussie, P., and Ferguson, K. M. (2005). Structural basis for inhibition of the epidermal growth factor receptor by cetuximab. Cancer Cell *7*, 301-311.

MacCallum, R. M., Martin, A.C., & Thornton, J.M. (1996). Antibody-antigen interactions: contact analysis and binding site topography. J Mol Biol *262*, 732-745.

McCoy, A. J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J Appl Cryst *40*, 658-674.

Medico, E., Russo, M., Picco, G., Cancelliere, C., Valtorta, E., Corti, G., Buscarino, M., Isella, C., Lamba, S., Martinoglio, B., Veronese, S., Siena, S., Sartore-Bianchi, A., Beccuti, M., Mottolese, M., Linnebacher, M., Cordero, F., Di Nicolantonio, F., & Bardelli, A. (2015). The molecular landscape of colorectal cancer cell lines unveils clinically actionable kinase targets. Nature Communications *6*, 7002.

Molina, M. A., et al. (2001). Trastuzumab (herceptin), a humanized anti-HER2 receptor monoclonal antibody, inhibits basal and activated HER2 ectodomain cleavage in breast cancer cells. Cancer Research *61*, 4744-4749.

Montagut, C., Dalmases, A., Bellosillo, B., Crespo, M., Pairet, S., Iglesias, M., Salido, M., Gallen, M., Marsters, S., Tsai, S. P.*, et al.* (2012). Identification of a mutation in the extracellular domain of the Epidermal Growth Factor Receptor conferring cetuximab resistance in colorectal cancer. Nat Med *18*, 221-223.

Morelli, M. P., Overman, M.J., Dasari, A., Kazmi, S.M.A., Vilar Sanchez, E., Eng, C., et al. (2013). Heterogeneity of acquired KRAS and EGFR mutations in colorectal cancer patients treated with anti-EGFR monoclonal antibodies. J Clin Oncol *31*.

Morin, A., Eisenbraun, B., Key, J., Sanschagrin, P.C., Timony, M.A., Ottaviano, M., & Sliz, P. (2013). Collaboration gets the most out of software. Elife *2*.

Newhall, K., Price, T., Peeters, M., Kim, T.W., Li, J., Cascinu, S., et al. (2014). Frequency of S492R mutations in the epidermal growth factor receptor: analysis of plasma DNA from metastatic colorectal cancer patients treated with panitumumab or cetuximab. Ann Oncol *25*, 105-117.

Ohno, S., Mori, N., & Matsunaga, T. (1985). Antigen-binding specificities of antibodies are primarily determined by seven residues of VH. Proc Natl Acad Sci USA *82*, 2945-2949.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. In Macromolecular Crystallography, C.W. Carter Jr., and R.M. Sweet, eds. (New York: Academic Press), pp. 307-326.

Pedersen, M. W., Jacobsen, H.J., Koefoed, K., Hey, A., Pyke, C., Haurum, J.S., Kragh, M. (2010). Sym004: a novel synergistic anti-epigermal growth factor receptor antibody mixture with superior efficacy. Cancer Research *70*, 588-597, doi: 510.1158/0008-5472.CAN-1109-1417.

Price, T. J., et al. (2015). Prevalence and outcomes of patients with EGFR S492R ectodomain mutations in ASPECCT: Panitumumab vs. cetuximab in patients with chemorefractory wild-type KRAS exon 2 metastatic colorectal cancer (mCRC). J Clin Oncol *33*, suppl. 3, abstract 740.

Schneider-Merck, T., et al. (2010). Human IgG2 antibodies against epidermal growth factor receptor effectively trigger antibody-dependent cellular cytotoxicity but, in contrast to IgG1,

only by cells of myeloid lineage. Journal of Immunology *184*, 512-520, doi: 510.4049/jimmunol.0900847.

Siravegna, G., Mussolin, B., Buscarino, M., Corti, G., Cassingena, A., Crisafulli, G., Ponzetti, A., Cremolini, C., Amatu, A., Lauricella, C., Lamba, S., Hobor, S., Avallone, A., Veltorta, E., Rospo, G., Medico, E., Motta, V., Antoniotti, C., Tatangelo, F., Bellosillo, F., Veronese, S., Budillon, A., Montagut, C., Racca, P., Marsoni, S., Falcone, A., Corcoran, R.B., Di Nicolantonio, F., Loupakis, F., Siena, S., Sartore-Bianchi, A., Bardelli, A. (2015). Clonal evolution and resistance to EGFR blockade in the blood of colorectal cancer patients. Nature Medicine *21*, 795-801, doi:710.1038/nm.3870.

Spector, N. L., & Blackwell, K.L. (2009). Understanding the mechanisms behind trastuzumab therapy for human epidermal growth factor receptor 2-positive breast cancer. J Clin Oncol *27*, 5838-5847.

Weitzner, B. D., Dunbrack, R.L., Jr., & Gray, J.J. (2015). The origin of CDR H3 structural diversity. Structure *23*, 302-311.

Xu, H., Schmidt, A.G., O'Donnell, T., Therkelsen, M.D., Kepler, T.B., Moody, M.A., Haynes, B.F., Liao, H.X., Harrison, S.C., & Shaw, D.E. (2015). Key mutations stabilize antigen-binding conformation during affinity maturation of a broadly neutralizing influenza antibody lineage. Proteins *83*, 771-780.

Yonesaka, K., Zejnullahu, K., Okamotom I., Satoh, T., Cappuzzo, G., Souglakos, J., et al. (2011). Activation of ERBB2 signaling causes resistance to the EGFR-directed therapuetic antibody cetuximab. Sci Transl Med *3*, 99ra86.

CHAPTER FIVE

Structural basis of specific targeting of oncogenic EGFRvIII by a camelid antibody/VHH domain.

SUMMARY

Despite its clinical significance, it is not understood how oncogenic EGFR variant III, an aberrant EGFR gene rearrangement, found in ~30-50% of glioblastoma multiforme (GBM) cases, causes documented constitutive transactivation of EGFR and MET tyrosine kinases. Because the EGFRvIII ectodomain mutation renders a unique tumor-specific antigen, it has been an attractive target antigen for several types of immunotherapies. Consequently, there has been significant interest in raising antibodies with specificity for EGFRvIII over wild-type EGFR. I present a 2.9 angstrom X-ray crystal structure of the extracellular region of the EGFR mutant vIII (sEGFRvIII) with an unpaired cysteine mutated to serine (C16S) in complex with a camelid antibody. This VHH domain, the variable domain from a heavy chain only camelid antibody, was generated by Paul van Bergen en Henegouwen (Utrecht University) by selection from a library for VHH domains that selectively bind EGFRvIII and not EGFR. I show that this VHH domain has a 25-fold specificity for sEGFRvIII over wildtype sEGFR. My X-ray crystal structure of the VHH in complex with the extracellular region of EGFRvIII reveals that this heavy chain only camelid antibody targets a novel epitope on EGFR domain IV that is occluded by the intramolecular 'tether' in wild-type EGFR. The structural basis for EGFRvIII targeting by this VHH domain is the first structural view of specific antibody targeting of EGFRvIII and supports further analysis of this VHH domain as a potential tool for specifically targeting this important oncogenic EGFR variant. As previously observed by others in the Ferguson laboratory, domain II of EGFRvIII in the sEGFRvIII/VHH structure is poorly ordered. Small angle X-ray scattering analysis suggest that domain II of sEGFRvIII adopts multiple conformations, I suggest may play a role in the ability of EGFRvIII to transactivate other receptor tyrosine kinases, such as MET.

INTRODUCTION

Mutation and overexpression of receptor tyrosine kinases (RTKs) is often associated with cellular transformation and is found in many human cancers. EGFR amplification and subsequent overexpression is a common molecular feature of glioblastoma multiforme (GBM), occurring at a frequency of 34-63% (Ekstrand, 1991; Heimberger, 2005; Libermann, 1985; Shinojima, 2003). Historically, one of the first viral oncogenes to be discovered from avian erythroblastosis virus was v-ErbB, which encodes a large extracellular truncation of the epidermal growth factor receptor (EGFR), and results in its elevated, constitutive tyrosine kinase activity and thereby cellular transformation (Downward, 1984). The most common EGFR mutation found in glioblastoma multiforme (GBM) is a similar extracellular truncation of EGFR called EGFR variant III (EGFRvIII) (Wong, 1992). EGFRvIII is reported to be present in ~30-50 % of glioblastoma cases and also less commonly in other cancers, such as lung, head and neck, and colorectal carcinomas (Gan, 2013).

EGFRvIII is a transforming, oncogenic alternative splice variant of EGFR that results in an exclusion of its exons 2-7 (it is also referred to as del2-7), which encodes a receptor missing a large portion of its ectodomain (mature amino acid residues 6-273) that includes all of domain I and much of domain II. The deletion includes both the domain I contribution to the ligand binding site and the part of domain II indispensable for wild-type EGFR dimerization (Dawson, 2005). EGFRvIII retains low-level, constitutive kinase activity and is classified as an oncogene (Gan, 2013). In this way it is reminiscent of v-erbB, which also has a large ectodomain deletion that includes the ligand-binding domains, leaving only a short portion of domain IV. The expression of EGFRvIII has also been recently reported to

regulated by epigenetic mechanisms, as inhibition of histone deacetylation can decrease expression of EGFRvIII in glioma cell lines by 50-80% (Del Vecchio, 2013).

Recent phosphoproteomic analysis of the glioblastoma cell lines U87MG and U373MG by White and colleagues has revealed EGFRvIII signaling networks that are quite distinct from wild-type EGFR, and involve transactivation the receptor tyrosine kinase MET (Huang, 2007; Johnson, 2012). Indeed, dual inhibition of EGFR and MET increases cytotoxicity of EGFRvIII-expressing glioma cell lines and greater survival in an intracranial mouse model of glioma (Greenall, 2015b; Huang, 2007). Moreover, transactivation of MET by EGFRvIII leads to resistance to the antibody rilotumumab, which binds to and neutralizes the MET ligand HGF (Pillay, 2009). It has also been found that ErbB2 can inhibit phosphorylation of EGFRvIII through possible heterodimerization with EGFRvIII (O'Rourke, 1998). In recent years it has also been suggested that constitutive phosphorylation of EGFRvIII and downstream signaling may require coexpression of EGFRvIII with wild-type EGFR (Fan, 2013), which would support the notion that hetero-oligomerization with other receptors is likely relevant for EGFRvIII-mediated signaling (Greenall, 2015b). It is not understood mechanistically how this cooperation of EGFRvIII and EGFR leads to transactivation of EGFR and MET tyrosine kinases and subsequent activation of STAT3/5, and thereby cellular transformation.

The wild-type EGFR extracellular region contains two-cysteine-rich domains, domains II and IV. The 50 ectodomain cysteine residues are organized into 25 disulfide bonds (Ferguson et al., 2003). The EGFRvIII deletion reduces the number of cysteine residues to 29 and leaves an unpaired cysteine in domain II (Cys283 in mature numbering for wild-type sEGFR). It is clear from other receptor tyrosine kinase mutations that the presence of an unpaired ectodomain cysteine residue often produce significant changes in receptor activation mechanisms and has the potential to alter receptor trafficking, signaling, and

proliferative capacity (Greenall, 2015a). For EGFRvIII, the presence of this unpaired cysteine was shown previously to contribute to disulfide mediated dimerization, and that mutation of the unpaired cysteine to a serine reduces dimerization and activation of EGFRvIII (Ymer, 2011). In addition, a xenograft model of EGFRvIII C16S showed reduced tumorigenicity compared to 'wild-type' EGFRvIII (Ymer, 2011). However, EGFRvIII transactivation of MET does not require covalent disulfide mediated dimerization, suggesting that intrinsic structural properties of EGFRvIII contributes to its hetero-oligomerization with other RTK's and its oncogenic signaling properties (Greenall, 2015b).

Previous unpublished work in the Ferguson laboratory suggests that a secreted protein with the unique N-terminal cysteine substituted with serine does not aggregate and is a good model protein to study the solution properties of sEGFRvIII. Furthermore, this previous work in collaboration with Graham Carpenter at Vanderbilt University structurally and biochemically examined EGFRvIII to determine how the unpaired cysteine residue in domain II contributes to intracellular trafficking of EGFRvIII. This work found that the unpaired cysteine of EGFRvIII contributes to disulfide-mediated crosslinking of EGFRvIII, which sequesters a population of EGFRvIII in the endoplasmic reticulum, consistent with a recent study (Greenall, 2015b). Mutation of this cysteine to serine permits trafficking of EGFRvIII to the cell surface and increases survival of an intracranial murine . Furthermore, in an X-ray crystal structure of the extracellular region of EGFRvIII (sEGFRvIII) C16S alone (K. R. Schmitz, K. M. Ferguson, unpublished data), the portion of domain II that remains in EGFRvIII is poorly ordered.

I observe by small-angle X-ray scattering that there is an increase in the radius of gyration for sEGFRvIIIC16S compared to 'tethered' EGFRvIII models. I speculate that this may be due to disorder in domain II and that this dynamic molecular feature may have

implications for understanding the structural basis of EGFRvIII-mediated hetero-oligomerization and oncogenic signaling.

Because EGFRvIII is a tumor specific antigen, there has been a lot of interest over the last two decades in raising antibodies that specifically target EGFRvIII over wild-type EGFR (Johns, 2002; Luwor, 2001; Wikstrand, 1995) for clinical applications. Many such antibodies are either in clinical development, such as mAb806 (Johns, 2002; Reilly, 2015), or used in the context of chimeric antigen receptor-T cell therapies for glioblastoma multiforme, such as 3C10 (Johnson, 2015; Takasu, 2003). I determined a 2.9 Å X ray crystal structure of the EGFRvIII C16S ectodomain in complex with VHH domain 34E5, that was selected on the basis of its ability to bind to EGFRvIII and not to wild type EGFR (P. van Bergen en Henegouwen, personal communication). I show that this VHH also has ~25-fold specificity for sEGFRvIII over wild-type sEGFR in a Biacore binding assay (Figure 5.5). VHH domains are increasingly emerging as potential tools for several therapeutic and imaging applications, and we previously reported X-ray crystal structures of three inhibitory nanobodies/VHH domains in complex with sEGFR (Schmitz, 2013). I discovered in the X-ray crystal structure of 34E5 bound to the extracellular region of EGFRvIII that the VHH domain targets a novel epitope on EGFR domain IV that in wild-type EGFR is sterically occluded by the intramolecular 'tether' between domains II and IV. Therefore, the VHH functionally mimics the intramolecular 'tether' between cysteine rich domains II and IV. Our structural analysis of nanobody/VHH targeting of a sterically occluded epitope in EGFR reveals a mechanism by which antibody specificity for oncogenic EGFRvIII can be achieved, and provides direct evidence of dynamic uncoupling of the intramolecular 'tether' in wild-type EGFR.

In the next sections I will describe unpublished data from others in the Ferguson lab that set the stage for my studies on sEGFRvIII.

Role of Cys16 in EGFRvIII Dimerization.

A soluble recombinant form of the extracellular region of EGFRvIII (sEGFRvIII) was previously produced using baculovirus infected Sf9 cells and purified exactly as described (Ferguson *et al.*, 2000; Ferguson *et al.*, 2003). It was shown that the protein secreted from the Sf9 cells existed in two distinct populations that could be partially separated during purification. One population was largely monomeric, as assessed by size exclusion chromatography. This population eluted from the initial Ni-NTA affinity purification column with low imidazole concentrations (50 - 75 mM). The second population eluted predominantly at higher imidazole concentrations (> 100mM) and showed clear evidence of aggregation by size exclusion chromatography (data not shown). Non-reducing SDS-PAGE suggested that this population contains multiple species of disulfide-linked oligomers. These two populations do not interconvert, further supporting that the aggregation is through covalent interactions.

It was reasoned that this aggregation could be due to the presence of an unpaired cysteine in the N-terminal region of sEGFRvIII. In contrast to the wild type EGFR, the ectodomain of the vIII mutations has an odd number of cysteines. The obvious choice for the unpaired cysteine is Cys16 in EGFRvIII (equivalent to Cys283 in mature EGFR, Fig. 5.1A. An sEGFRvIII variant with a serine at position 16 (C16S) was secreted from Sf9 cells as a largely monomeric and monodisperse protein. The majority of this protein behaved as a single monomeric species, although there remains a small (< 1%) population of disulfide-linked aggregates (Ferguson laboratory unpublished data, not shown).

To assess the influence of C16 in the oligomerization of EGFRvIII in cells, the Ferguson laboratory collaborated with Graham Carpenter's lab to express the full length wild-type EGFRvIII or EGFRvIII C16S mutant in HEK 293T cells and analyzed the expressed protein on non-reducing SDS gels. The Carpenter lab showed that EGFRvIII formed an

equivalent level of monomers and dimers, while the EGFRvIII C16S mutant formed only monomers.

Influence of C16S on EGFRvIII Intracellular Trafficking.

Previous reports have shown that the EGFRvIII is not trafficked efficiently to the cell surface, but it is not understood what molecular mechanisms contribute to this inefficient processing through the secretory pathway. A population of EGFRvIII receptors are sequestered in the secretory pathway (Wikstrand, 1997), and the population of EGFRvIII that does reach the cell surface are not internalized rapidly (Grandal, 2007). A previous Ferguson laboratory collaboration with Graham Carpenter's laboratory found a direct role of the C16S mutation on EGFRvIII trafficking. The Carpenter lab showed in these experiments that the vast majority of wild-type EGFR is present at the cell surface with a small intracellular pool detectable. In contrast, the results show that EGFRvIII is predominantly present in intracellular pools on a reticular lattice that resembles the distribution of ER. Similar experiments of EGFRvIII with the unpaired cysteine residue mutated to serine suggest that the intracellular pool of EGFRvIII C16S is diminished relative to EGFRvIII (G. Carpenter, personal communication, data not shown). These data in addition show that in HEK293T cells, EGFRvIII exists as a mixture of monomers and dimers. The ER localization of EGFRvIII is consistent with the slow conversion of monomers to dimer species due to the delayed formation of intermolecular compared to intramolecular disulfides, as observed with secreted immunoglobulins (Nag, 2015).

If EGFRvIII is predominately located in the endoplasmic reticulum, then its glycosylation would be of the immature or high-mannose type. In contrast, the EGFRvIII C16S mutant, if more efficiently processed to the cell surface, should have complex oligosaccharide chains. The types of oligosaccharide present on these molecules can be distinguished by sensitivity to endoglycosidase H, which selectively cleaves high-mannose oligosaccharides (Maley, 1989; Tarentino, 1974). The Carpenter lab demonstrated that the

while EGFRvIII is sensitive to Endo H digestion, the C16S is mutant is relatively resistant. The results of Endo H digestion of EGFRvIII compared to the completely non-glycosylated form of EGFRvIII produced in the presence of tunicamycin (Oslowski, 2011) further support the conclusion that the cysteine mutant contains complex glycans where the 'wild-type' EGFRvIII contains predominantly high mannose glycans (G. Carpenter, personal communication, data not shown). Therefore, while EGFRvIII is characterized by the presence of high-mannose oligosaccharides typical of ER localized proteins, the EGFRvIII C16S mutant contains complex chains typical of cell surface proteins.

X-ray crystal structure and small angle X-ray scattering of sEGFRvIIIC16S.

A previous member of the Ferguson laboratory, Karl Schmitz, determined a 2.5 Å resolution X ray crystal structure of sEGFRvIII C16S alone. Refinement of this structure revealed that the part of domain II that is left in EGFRvIII (mature amino acids 274-310) is disordered. Domains III and IV of the sEGFRvIII C16S structure are well ordered and adopt a very similar conformation to that observed in other structures of sEGFR (Fig. 5.2B). Domain II is not well ordered and very little of this part of the protein can be modeled, although fragmented density provide some hints as to the location of this key part of the sEGFRvIII molecule.

There is an antiparallel dimer in the asymmetric unit, with two disordered domain II units in apposing molecules facing toward one another. If this oligomer is populated at the cell surface, any unpaired cysteine in domain II of EGFRvIII (such as C16) would be poised to form an intermolecular disulfide bond. However, in the absence of disulfide crosslinking the sEGFRvIIIC16S is monomeric as assessed by sedimentation equilibrium analytical ultracentrifugation of 8 μM receptor (Figure 5.1). As noted above, even this form of sEGFRvIII with an even number of cysteines has a propensity to form disulfide mediated

oligomers when secreted from insect cells, suggesting that there may be some level of disulfide mixing in the domain II portion of sEGFRvIII. For the wild type EGFRvIII both when secreted from Sf9 cells and when expressed as the full-length receptors in HEK293T cells, there is clear indication that disulfide mediated dimers form.

RESULTS

Small-angle X-ray scattering analysis of sEGFRvIII C16S reveals a larger Rg compared to 'tethered' sEGFRvIII models.

I turned to small angle X ray scattering (SAXS) analysis of the purified EGFRvIII extracellular region – using the monodisperse C16S mutant, as the 'wild-type' EGFRvIII extracellular region is prone to aggregation – to investigate intrinsic structural properties of the EGFRvIII extracellular region in solution that may play a role in its transactivation of other RTK's and its internalization properties. I performed SAXS experiments at concentrations ranging from ~2.5 mg/ml to ~12 mg/ml. Linear Guiner regions were observed at all three concentrations, suggesting no sign of aggregation. Analysis of the pairwise interatomic distance distribution functions calculated by inverse Fourier transformation of the raw scattering data revealed a ~5 Å increase in the radius of gyration (R_g) of EGFRvIII C16S compared to either models of 'tethered' EGFRvIII (based on PDB ID 1NQL) or the model of EGFRvIII C16S alone determined by Karl Schmitz. The P(r) interatomic vector distribution is overall shifted towards larger distances, but with about the same maximum interatomic distance (D_{max}) of about 100 Å, consistent with the D_{max} calculated for a model of 'tethered' EGFRvIII alone (97 Å) in CRYSOL (Svergun, 1995). These data are consistent with the ability of domain II being able to adopt alternate conformations and the increased flexibility of

the domain III/IV relationship as observed by the altered hinge angles between domains III and IV in the X-ray crystal structures of sEGFRvIII. I speculate that such increased flexibility of domain II as a result of a loss of stabilizing interactions with domain I could allow EGFRvIII to more accessibly heterodimerize with and cause transactivation of other receptor tyrosine kinases and covalently dimerize with other EGFRvIII receptors.

X-ray crystal structure of sEGFRvIII C16S in complex with 34E5 VHH.

I had hoped to stabilize an alternate conformation of domain II to observe crystallographically by exploiting a camelid antibody that had been phage display selected for binding to EGFRvIII. The VHH library had been generated from the lymphocytes of *lama glama* that had been immunized with A431 cells (a breast cancer cell line that overexpresses EGFR). I purified and crystallized sEGFRvIII C16S in complex with this EGFRvIII-selected VHH domain (34E5). Crystals of P1 symmetry were produced by the hanging drop vapor diffusion method and diffracted to 2.9 Å resolution and the structure was solved by molecular replacement by sequentially identifying domain III, IV, and subsequently the VHH.

The X ray crystal structure of the VHH/sEGFRvIII complex reveals that this nanobody/VHH targets a novel epitope on domain IV that is sterically occluded by the intramolecular 'tether' between domains II and IV observed in wild-type sEGFR structures (Figure 5.6). Similar 'tethers' are observed in structures of ErbB3, and ErbB4. In contrast, ErbB2 is constitutive 'extended' as a monomer (Lemmon, 2014). Therefore, this VHH gains specificity for EGFRvIII through targeting a sterically occluded epitope in wild-type EGFR that is not sterically occluded in EGFRvIII due to its lack of a 'tether'.

The conformation of domain III is very similar to that in other sEGFR structures. Domain III in the sEGFRvIII C16S/VHH complex overlays with domain III in the

sEGFR/cetuximab complex structure (PDB ID 1YY9) with an RMSD of 0.439 Å and with domain III in 1NQL with an RMSD of 0.446 Å. Domain IV exhibits a larger ~10 degree 'swing' out from sEGFR in PDB ID 1YY9 to adopt a more extended conformation. This domain IV trajectory is ~15 degrees in the opposite direction from the antibody unbound structure of sEGFRvIII (Figure 5.5). The interface area between the VHH and EGFRvIII is 766 $Å^2$, with a shape complementarity statistic of 0.674 (Lawrence, 1993), comparable to therapeutic monoclonal antibodies that target EGFR (Li, 2008; Li et al., 2005).

The majority of the VHH interaction with EGFRvIII domain IV is made by contacts from CDR3, which extends into a groove in domain IV between two strands (Figure 5.4C). At the focal point in CDR3, there is a cluster of polar interactions with the region of EGFRvIII domain IV that normally interacts with the β hairpin in domain II in tethered sEGFR structures (Figure 5.4C). At the tip of this loop, Arg102 in the VHH forms hydrogen bonds with the mainchain backbone carbonyls from Ala559 and Cys558 in EGFRvIII domain IV. The guanidinium moiety of this arginine sidechain is also stabilized in this position by a cation-pi interaction with its neighboring residue Tyr101, which also forms several polar mainchain interactions with domain IV, as well as an aromatic pi-stacking interaction with the sidechain of His591. Along with another flanking tyrosine, Tyr104 in CDR3, and Tyr37 from CDR1, these tyrosines form important van der Waals interactions with EGFRvIII. An ordered lysine, K96 in CDR3 hydrogen bonds with the backbone carbonyl from Pro565 in EGFRvIII. In addition, Gln557 of EGFRvIII hydrogen bonds with the backbone carbonyl of Arg102 in the VHH. His594 in EGFRvIII in engaged in a bivalent polar interaction with Asp54 in CDR2 and a possible cation-pi interaction with Phe31 in CDR1. Finally, Ser56 in CDR2 is engaged in a hydrogen bond with an ordered Lys585 in EGFRvIII.

More of domain II can be resolved in this structure than in the antibody-unbound structure, possibly due to a weak crystal contact with the portion of the VHH distal to its

CDR's. In the antibody unbound structure, arginine 44/310 at the domain II/III boundary makes a salt bridge interaction with glutamate 110/376 in domain III. This interaction is also observed in other tethered sEGFR structures but is broken in the ligand bound structures due to dramatic local conformational change at the domain II/III boundary that must occur for EGFR to transitions from the tethered unliganded state to the extended ligand bound conformation (Fig. 5.2G). The conformation of sEGFR with the last disulfide-bonded module of domain II (m8) tucked "under" domain III must be favored even in the absence of the domain II/IV tether or of an intact domain I/II fragment. Consistent with this the isotropic temperature factors in the domain II/III boundary region of tethered sEGFR are low whereas for the ligand bound structures this region has higher than average B-values.

In the VHH bound structure, electron density for the Arg310 sidechain is clearly resolved in one of the two complexes in the asymmetric unit and absent in the other. In the complex with density for R310, it is hydrogen bonding with a Ser117 in the VHH, in a crystal lattice contact. Glu376 is instead engaged in a salt bridge with Arg403 in domain III, in both complexes in the asymmetric unit. Electron density N-terminal to the first disulfide module in domain II remains too poorly resolved to accurately model. Isotropic B factors for the part of domain II that is modeled are quite high (>120), suggesting a high degree of disorder and possible flexibility in this portion of EGFRvIII.

DISCUSSION

Previous work in the Ferguson laboratory showed that the unpaired cysteine in EGFRvIII contributes to disulfide-mediated dimerization of EGFRvIII, consistent with a previous study that the free cysteine contributes to oligomerization of EGFRvIII (Greenall, 2015b). This effect contributes to EGFRvIII being sequestered in early in the secretory

pathway, in the endoplasmic reticulum. While abrogating this effect is not required for MET transactivation, it does cause decreased oncogenic potential and greater survival of a murine intracranial model of glioma (Ymer, 2011). This suggests that sequestration of EGFRvIII in the endoplasmic reticulum affects its downstream signaling – suggesting that the intracellular location of these mutated receptors is critical for their capacity to cause oncogenic signaling. There are several well-documented examples of receptor tyrosine kinases having distinct and oncogenic signaling properties when their trafficking is and processing in the secretory pathway is altered (Choudhary, 2009; Schmidt-Arras, 2009). One related example of this is TrkAIII, another oncogenic alternative splice variant of the RTK TrkA that removes its exons 6 and 7, resulting in a removal of its D4 Ig-like domain and causing its mislocalization, ERGIC/COPI-associated activation, and oncogenic Akt signaling (Farina, 2015). Two examples of internal tandem duplication (ITD) events that cause sequestration of RTKs in the endoplasmic reticulum or Golgi apparatus also include FLT3 and KIT (Schmidt-Arras, 2009). The c-KIT ITD in its juxtamembrane region, found in childhood acute myeloid leukemia (AML), has been linked to ligand-independent KIT activation and transformation of a Ba/F3 model (Corbacioglu, 2006).

My solution X-ray scattering analysis of the in vitro purified cysteine to serine sEGFRvIII mutant further reveals that domain II flexibility may also be a factor in causing a decreased energetic barrier for homo- or hetero-dimerization of EGFR, consistent with our hypothesis for extracellular oncogenic point mutations in the EGFR (Bessman, 2014). This dynamic molecular feature of EGFRvIII may result from the loss of autoinhibitory interactions from the domain I/II interface that may function normally to stabilize and regulate domain II conformation. Removal of these interactions, as well as N-linked glycosylation sites in domain I, may cause an increase in the degrees of freedom in domain II, consistent with its low X-ray scattering properties in two X-ray crystal structures. Some reports have argued that EGFRvIII-mediated oligomerization relies on entirely on covalent disulfide mediated

oligomerization. Previous work in the Ferguson laboratory in collaboration with the Carpenter laboratory found that covalent dimerization of EGFRvIII through disulfide mixing of cysteines in domain II likely contributes to its sequestration and primary localization in the endoplasmic reticulum, and hence affecting EGFRvIII signaling. However, I speculate that instrinsic structural flexibility of EGFRvIII ectodomain also may contribute to its ability to transactivate other receptor tyrosine kinases, including EGFR and MET.

My structural analysis of the 34E5 VHH/sEGFRvIII complex reveals a unique mode of antibody specificity for an oncogenic EGFR variant, and is the first structure of EGFRvIII bound to an antibody with engineered specificity, of which there are several, including 3C10 and mAb806 (Johns, 2002; Luwor, 2001; Reilly, 2015). The novel epitope of this VHH with engineered specificity for EGFRvIII for the first time reveals a mechanism by which antibody specificity for an oncogenic EGFR variant specificity can be achieved through exploitation of the presence of the intramolecular 'tethering' interaction between cysteine rich domains II and IV. Moreover, because this epitope is only accessible in wild-type sEGFR when this tether is transiently broken, the energetic difference, in terms of Gibbs free energy, between 34E5 binding to wild-type EGFR and EGFRvIII yields a thermodynamic estimate for the energetics of 'tether' breaking of ~1.7 kcal/mol. This is the first direct energetic view of the thermodynamics of equilibrium between 'tether' formation and disruption.

Antibodies that have been raised to the ErbB/HER family tend to bind to recurring epitopes on EGFR. Cetuximab, panitumumab, and necitumumab, for example, all bind to the same region of domain III, and block ligand binding to EGFR, along with two inhibitory VHH domains (Li, 2008; Li, 2005; Schmitz, 2009; Schmitz, 2013). One inhibitory anti-EGFR VHH domain and KTN3379, an anti-ErB3 antibody that is currently in phase II trials (Kolltan Pharmaceuticals, New Haven, CT), bind to the domain II/III interface and lock the conformational transition to the activated dimeric state (Lee, 2015; Schmitz, 2013). Trastuzumab/Herceptin^{TM}, an anti-ErbB2/HER2 antibody, targets domain IV of ErbB2/HER2,

utilizing a similar utilizing a similar binding mode as the 34E5 VHH for specifically recognizing EGFRvIII (Figure 5.6) (Cho, 2003). Pertuzumab/Perjeta™, another anti-ErbB2 antibody, interacts with the β hairpin in domain II that also forms the intramolecular 'tethering' interaction in the other ErbB receptors (Franklin, 2004).

One possible mechanism that may contribute to the recurrence of similar binding modes for antibodies that have been selected for binding to specific EGFR variants is a 'chaperone' effect by glycans that surround these epitopes. For instance, N328, N389, and N420 all 'straddle' the shared cetuximab, panitumumab, and necitumumab epitope on domain III. In the case of the 34E5 VHH domain, there are two N-linked glycosylation sites that may also serve a similar function on N544 and N579. It is widely established that N-linked glycosylation plays an important role in both proper folding of the EGFR (Gamou, 1988; Gamou, 1989), ligand binding (Tsuda, 2000), and membrane interactions (Kaszuba, 2015). In addition, complexity of N-linked glycans, including sialylation and fucosylation, has recently been reported to play a role in regulating dimerization and activation of EGFR in lung cancer cell lines (Liu, 2011). I suggest that N-linked glycans may also play a role in directing anti-ErbB/HER antibodies to their epitopes.

My discovery of this novel epitope for the EGFRvIII-specific VHH corroborates the utility of the intramolecular 'tether' interaction in ErbB receptors can be exploited for specific antibody targeting of oncogenic EGFR variants lacking this 'tether', much like trastuzumab and pertuzumab target ErbB2/HER2. As the buried surface area of the tether is relatively small, the smaller VHH paratope is uniquely suited to gain specificity for EGFRvIII through this mechanism. The 25-fold specificity gained for EGFRvIII through this mechanism could potentially be exploited to target novel immunotherapies that target EGFRvIII, and suggests other possible mechanisms for gaining specificity for ErbB2/HER2 specific antibodies. Similar specific recognition of oncogenic forms of RTKs may offer important avenues for specific targeting of tumor specific antigens by cancer immunotherapies.

METHODS

Protein production.

A pFastbac vector for generation of sEGFRvIII was created previously by others in the Ferguson laboratory by replacing codons 6 – 273 of wild-type sEGFR (Sugawa et al., 1990) with a single glycine codon using standard molecular biology methods. The plasmid to create the cysteine to serine variant of sEGFRvIII was created using standard PCR mutagenesis methods. Both proteins were expressed and purified as secreted proteins in baculovirus-infected Sf9 cells exactly as previously described for sEGFR (Ferguson et al., 2000; Li et al., 2005). The 34E5 VHH was subcloned into a pET22b bacterial expression vector with a C-termal Myc tag followed by a hexahistidine tag, and an N-terminal pelB leader sequence to direct periplasmic expression. This construct was provided generously by the laboratory of Dr. Paul van Bergen en Henegouwen (Utrecht University, Utrecht, Netherlands). 34E5 was expressed and purified from BL21(DE3) *E. coli* cells expressing pLysS using autoinduction (Studier, 2005) at 20 degrees for 20 hours. Periplasms were lysed by one round of freeze-thaw of the cell pellet, in 25 mM sodium phosphate 150 mM NaCl pH 7.8 10% glycerol. The VHH was purified by Ni-NTA chromatography followed by size exclusion chromatography using a Superose 12 column in 25 mM HEPES 150 mM NaCl pH 8, and was >95% pure by Coomassie staining.

Surface Plasmon Resonance

Surface Plasmon Resonance (SPR) experiments were carried out on a Biacore 3000 instrument exactly as we have previously described (Ferguson *et al.*, 2000). In brief, EGF

(purchased from Chemicon, Inc.) in 10 mM sodium acetate pH 4 was immobilized an activated CM5 surface. A series of samples of the relevant sEGFR protein (in 25 mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.005% nonidet-P20, pH 8.0) of different concentrations were injected over this surface at a flow rate of 5 μl/min with a contact time of 15 minutes, and a dissociation time of 1000 sec., with a 1 minute pulse of regeneration with 10 mM sodium acetate pH 4.5 1 M NaCl. Multiple rounds of regeneration did not impair sEGFR binding. The VHH surfaces were not stable for over a week, and therefore each independent experiment was done on a freshly prepared surface. Binding experiments were performed independently three times. The equilibrium SPR responses for each sample plotted against the sEGFR concentration, and normalized against the Bmax fitted for sEGFRvIII (applying a normalization factor of 1.45 to account for the mass difference between sEGFRvIII and sEGFR wild-type). Data were fit to a simple one-site Langmuir binding model using Prism 6 (GraphPad Software, Inc., La Jolla, CA).

Small-angle X-ray scattering (SAXS)

Small angle X ray scattering analysis of sEGFRvIII C16S was performed in house using a Rigaku S-MAX3000 pinhole camera system, with a Rigaku 007HF rotating anode source and a Rigaku 300 mm wire grid ASM DTR 200 detector, with 60 minute exposures at 4°C. Scattering data were radially-averaged, normalized, and reduced to two-dimensional plots using the SAXSgui software, and intensity data from buffer exposures were then subtracted out. Radius of gyration (R_g) values were determined from Guinier plots using the Primus software package using a q range where $qR_g < 1.3$ (Konarev et al., 2003). The

scattering intensity at a scattering angle of 0 was also determined from Guinier analysis, and was plotted as a function of receptor concentration, to ensure scattering increased linearly with receptor concentration (Figure 5.2). The maximum interatomic distance (D_{max}) was obtained by examining P(r) curves generated through inverse Fourier transformation by an FFT algorithm in the Gnom software package (Svergun, 1992). Briefly, for each scattering dataset, P(r) curves were calculated for a range of D_{max} from 70 to 150 Å, in 5 Å increments. D_{max} was determined by identifying the value that gave the best fit to the experimental scattering data. P(r) curves for structural comparison were calculated from structural models by using the CRYSOL software package (Svergun, 1995).

Crystallization, data collection and structure determination

To purify the sEGFRvIII C16S/34E5 complex, a 2-fold molar excess of the VHH was incubated with sEGFRvIII C16S, and the complex purified by size exclusion chromatography using a Superose 6 column (GE Healthcare) in 25 mM HEPES pH 8 150 mM NaCl. The complex was concentrated to 70-90 μM and crystallized by hanging drop vapor diffusion at 20 degrees in 0.5 μl drops that contained a 1:2 v/v protein to reservoir ratio in 100 mM ammonium phosphate pH 7.1-7.6 15-20 % PEG3350. Large plate crystals were grown by either cracking the cover slips and sealing over the cracks with grease or by microseeding. Crystals appeared within 2 days and were stable for less than a week. Crystals were briefly cryoprotected in 100 mM ammonium phosphate pH 7.1 containing 20% PEG400 and 20% PEG3350. X-ray diffraction data to 2.9 angstrom resolution were collected at APS beamline 23-ID B using a MAR CCD detector. The data were processed in HKL2000 (Otwinowski and Minor, 1997). Data collection statistics are summarized in Table 1. The structure of sEGFRvIII was solved by the method of molecular replacement (MR) using the program PHASER (McCoy, 2007). As a search model the domains III and IV of wild-type EGFR (amino acids 310-480 and 481- 614 from PDB ID 1YY9, respectively) were first located

separately and subsequently together. Solutions for domain III were validated by identifying electron density for ordered N-acetylglucosamine residues (omitted from the search model) at N-linked glycosylation sites N328 and N389, and these solutions were fixed in subsequent searches.

The VHH was identified by using a sequence related camelid VHH (PDB ID 4NC2) with the complementarity determining regions (CDR's) omitted. The CDR regions were manually rebuilt using COOT (Emsley, 2010) after density modification and solvent flattening implemented in PHENIX AUTOBUILD (Terwilliger, 2008). Refinement was performed in PHENIX (Adams, 2010) and used iterative rounds of real space refinement with Ramachandran restraints followed by unrestrained reciprocal space refinement and model rebuilding.

Structure Analysis.

Structures were overlaid and RMSD values were calculated using the programs PYMOL (Delano Scientific, Palo Alto, CA), or COOT (Emsley, 2010), and reflect main chain atoms only. Shape complementarity values were calculated using the Sc module in CCP4 (Lawrence, 1993), excluding saccharide, ligand, and solvent molecules. Buried interface surface area was calculated using the PISA server. Most of the X ray diffraction and structural analyses were completed using software compiled on SBGrid (Morin, 2013).

CHAPTER FIVE TABLES

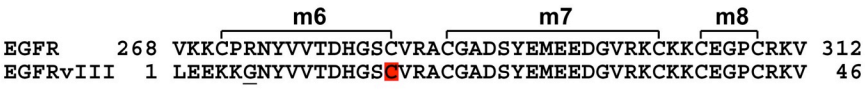
Table 5.1. Equilibrium binding constants of sEGFRvIII and sEGFRvIIIS16C binding to immobilized EGF, FabC225, and 34E5 VHH.

Immobilized Ligand	K_D values for binding of indicated analyte			
	sEGFR wild-type	sEGFRd3	sEGFRvIII (monomeric)	sEGFRvIIIC16S
EGF	130± 3 nM	1.7 ± 0.1 μM	2 ± 0.13 μM	2.2 ± 0.22 μM
FabC225	2.3± 0.5 nM	1.7 ± 0.1 nM	N.D.	1.1 ± 0.1 nM
34E5 VHH	668.3±107.8 nM	N.D.	35.6±11.3 nM	40.1±3.5 nM

. **Table 5.2. X-ray diffraction data collection and refinement statistics**

	sEGFRvIII C16S/34E5 VHH
Data collection	
Space group	P1
Cell dimensions	
a, *b*, *c* (Å)	66.5, 80.5, 92.2
α, β, γ (°)	107.04, 103.58, 104.56
Resolution (Å)	2.9 (3.0-2.89)
R_{merge}	0.087 (0.76)
$I / \sigma I$	23.7 (1.7)
Completeness (%)	96.6 (83.3)
Redundancy	3.8 (3.3)
Refinement	
Resolution (Å) (d_{min})	43-2.9 (2.97-2.89)
No. unique reflections	35,331
R_{work} / R_{free}	0.263/0.283 (0.433/0.509)
No. atoms	
Protein	6,694
Ligand/ion	11
Waters	
B-factors	
Protein	75.8
Ligand/ion	
Water	
R.m.s. deviations	
Bond lengths (Å)	0.002
Bond angles (°)	0.615
Ramachandran favored (%)	92.5
Ramachandran outliers (%)	0.7
Rotamer outliers (%)	0

A



B

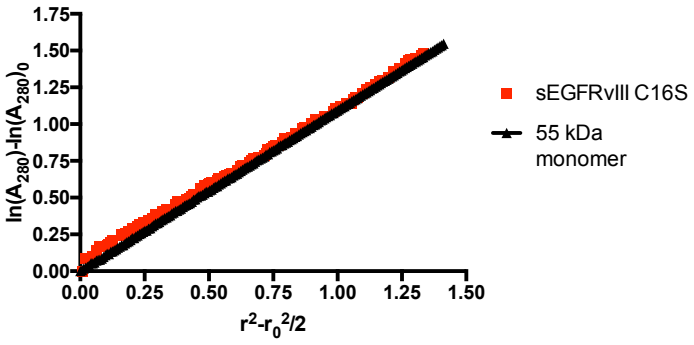


Fig. 5.1. Structural and biochemical properties of sEGFRvIII and sEGFRvIII C16S. (**A)** Sequence alignment of EGFRvIII with wild-type EGFR. Unpaired cysteine at position 16 in EGFRvIII is highlighted in red. **(B)** Sedimentation equilibrium analytical ultracentrifugation analysis demonstrates that EGFRvIII has a molecular mass of 55 kDa, corresponding to a glycosylated sEGFRvIII monomer (protein only molecular weight 41.1 kDa).

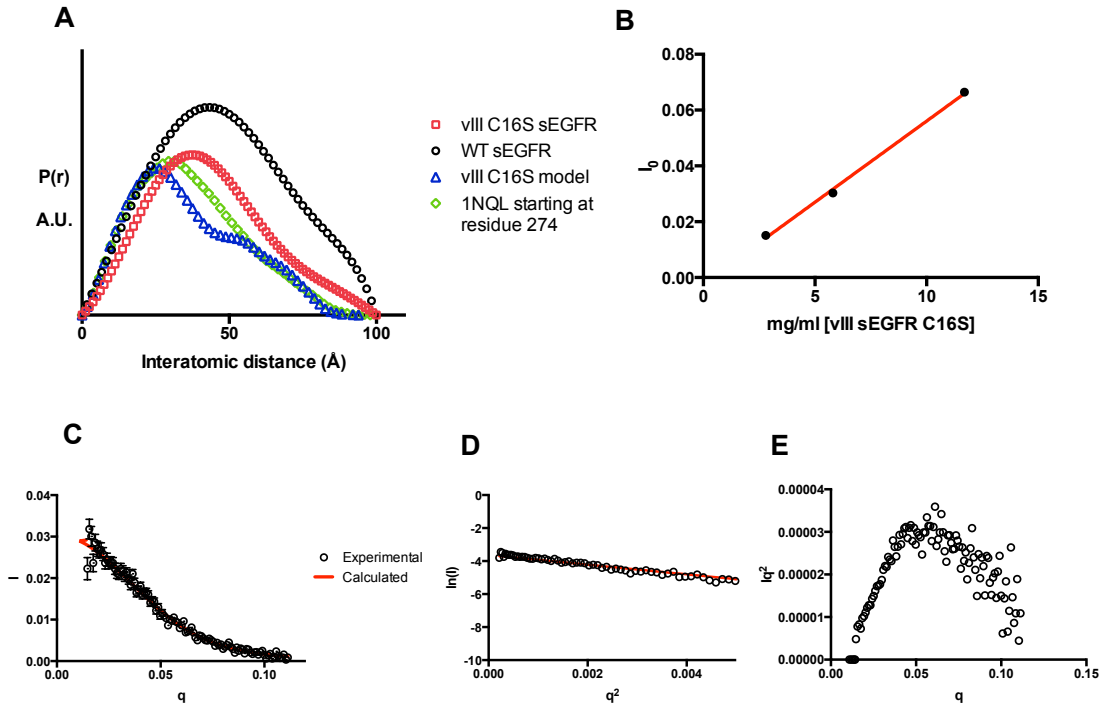


Fig. 5.2 Small angle X ray scattering analysis of EGFRvIII reveals domain II flexibility. **(A)** P(r) pairwise interatomic vector distribution functions calculated from solution X-ray scattering data of sEGFRvIII C16S (red) or calculated from indicated models. A 'tethered' EGFRvIII model was derived from PDB ID 1NQL (green). The experimentally determined sEGFRvIII C16S with a limited portion of domain II modeled is shown in blue. Reciprocal space Rg of sEGFRvIII C16S is 33.48, where that of a 'tethered' EGFRvIII model, derived from PDB ID 1NQL is 28.74. **(B)** Intensity at scattering angle 0 (I_0) plotted as a function of sEGFRvIII C16S concentration. **(C)** Raw scattering data plotted with maximum likelihood-derived errors in the fit to the calculated P(r), at a concentration of 6 mg/ml **(D)** Guinier region of the data shown in panel (C) shows no sign of protein aggregation under the

conditions of the experiments. **(E)** Kratky analysis of the data shown in (C) shows that the protein is globular and not denatured.

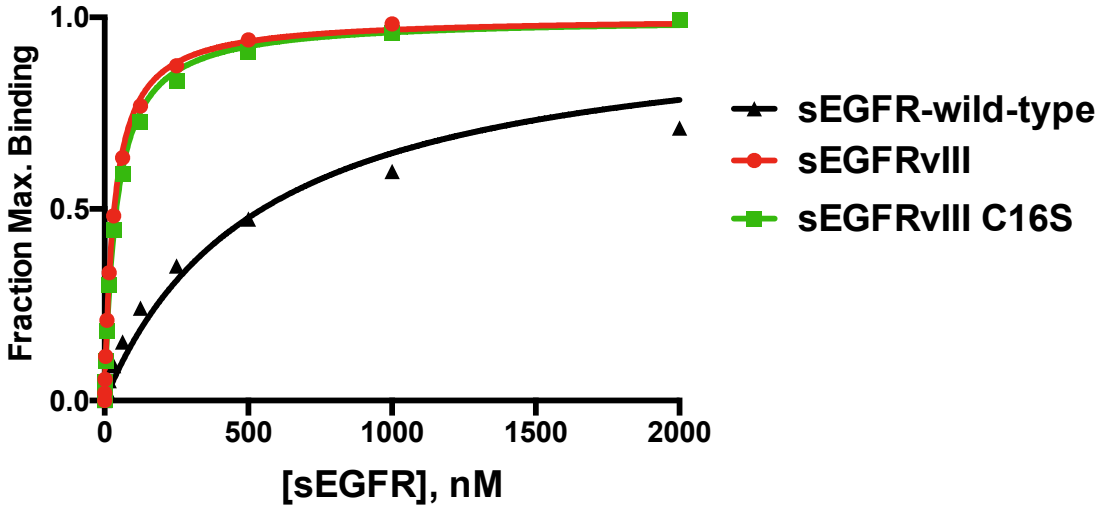


Figure 5.3 34E5 VHH has engineered specificity for EGFRvIII. Samples of the indicated concentrations of sEGFR and variants were passed over Biacore surfaces to which the VHH 34E5 had been immobilized. 34E5 samples were immobilized on Biacore CM5 chips using amine coupling. Binding assay and data processing were conducted as described in the experimental procedures. Representative, normalized binding curves are shown and mean K_D values are reported in Table 5.1.

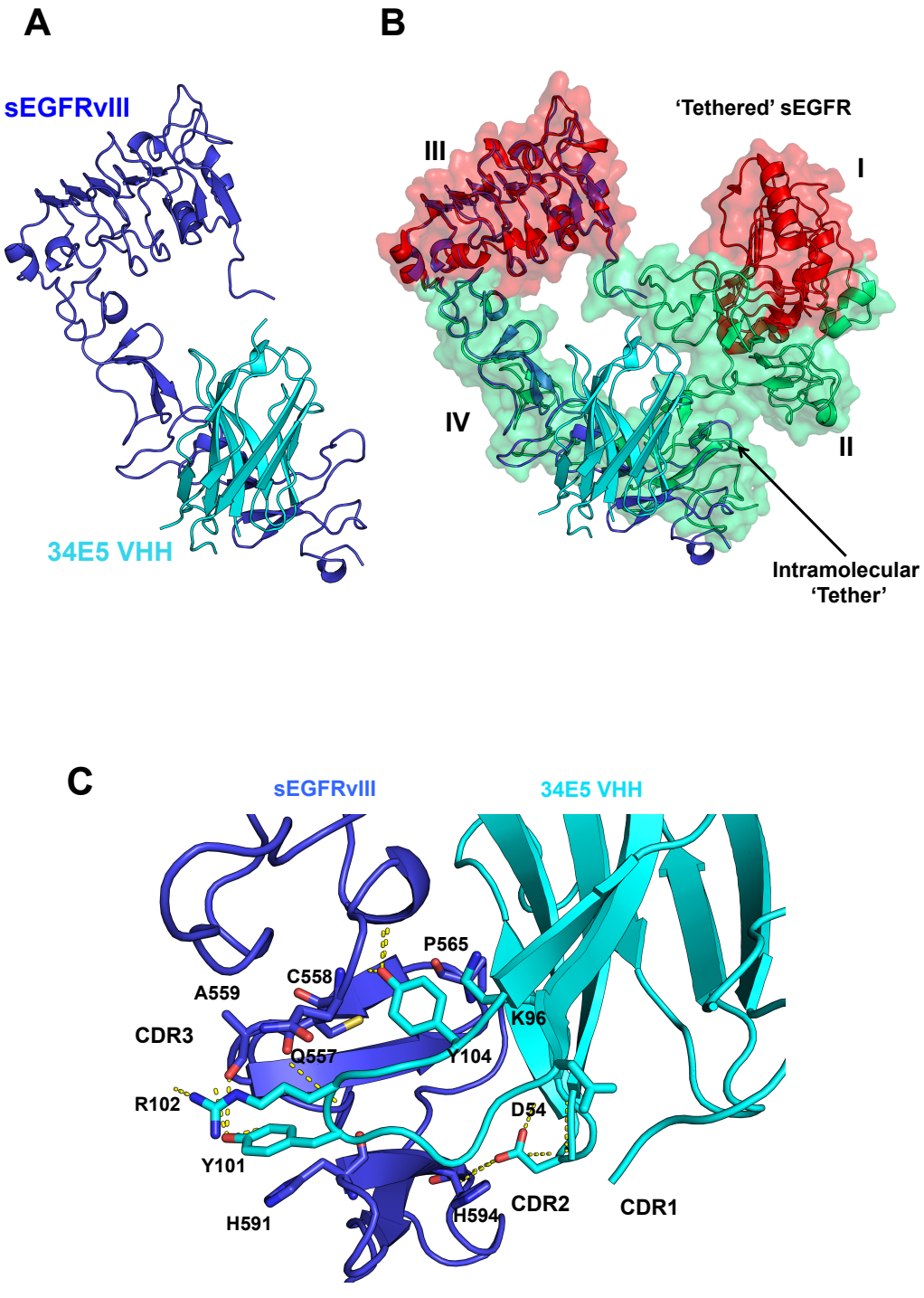

A
sEGFRvIII
34E5 VHH
B
'Tethered' sEGFR
III
I
IV
II
Intramolecular
'Tether'
C
sEGFRvIII
34E5 VHH
P565
C558
A559
K96
CDR3
Q557
Y104
R102
D54
Y101
H591
H594
CDR2
CDR1

Figure 5.4 Structural basis for specific targeting of EGFRvIII by 34E5 VHH. (A) The X-ray crystal structure of 34E5 VHH in complex with EGFRvIII reveals that 34E5 VHH binds to domain IV of EGFRvIII, to a region that is sterically occluded by the intramolecular 'tethering' interaction between domains II and IV, as shown in panel **(B). (C)** A detailed view of the molecular basis of 34E5 targeting of EGFRvIII. As shown, and discussed in the results section, the majority of the 34E5 interaction with EGFRvIII is driven by CDR3 hydrogen-bonding and van der Waals interactions, with some contributions from Asp54 in CDR2 and CDR1 (not shown).

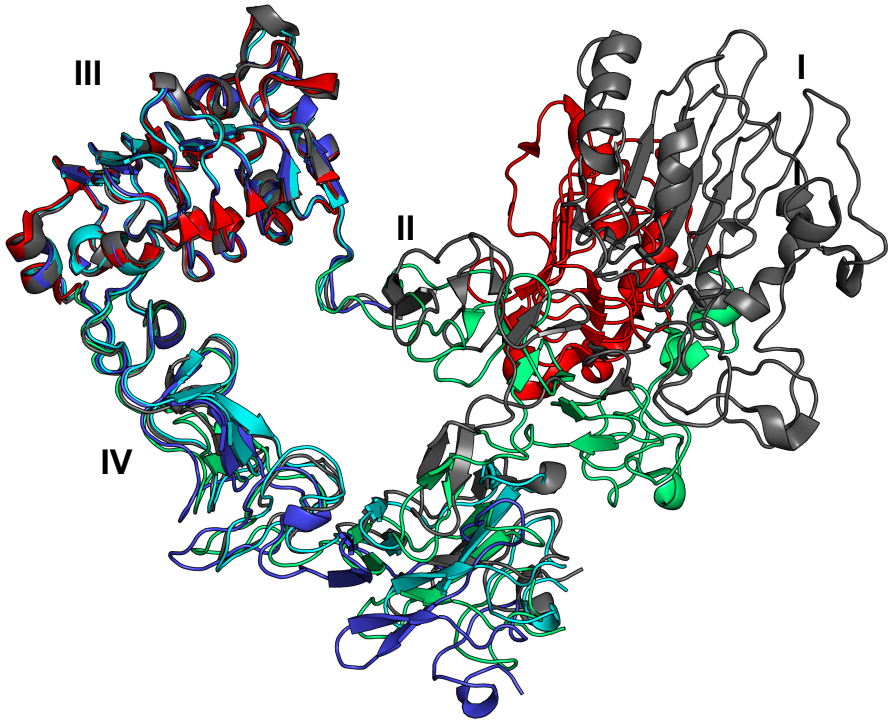


Figure 5.5 Structural comparison of sEGFR and sEGFRvIII C16S domain conformations. Structural overlay of domain III from known 'tethered' EGFR structures reveals distinct interdomain relationships. In red and green is the tethered structure determined at low pH with EGF (PDB ID 1NQL), in gray is the tethered structure observed in the cetuximab complex structure (PDB ID 1YY9). The blue and cyan models are of EGFRvIII. The cyan model is the structure of EGFRvIII C16S alone, determined previously. The blue model is of EGFRvIII C16S in complex with 34E5 VHH reported here. Comparison of domain IV conformation relative to domain III reveals a range of hinge angles between the two domains in tethered sEGFR models. The domain II/III interface also has two distinct geometries – one that is more like the inactive, low pH structure, and one that is more like the cetuximab complex structure. The part of domain II that can be modeled in the EGFRvIII

C16S structure in complex with the 34E5 VHH (blue) more closely resembles the cetuximab bound conformation of domain II (gray).

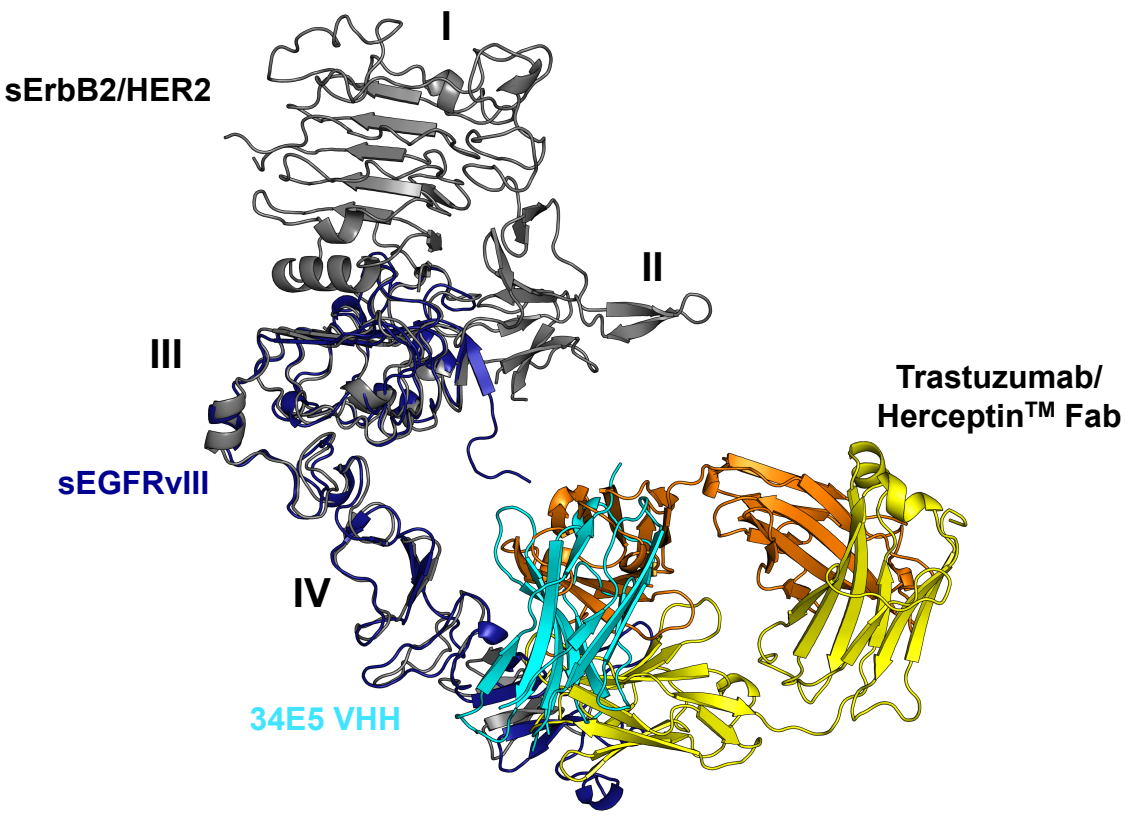


Figure 5.6 34E5 VHH utilizes a similar binding mode to specifically target EGFRvIII as trastuzumab uses to target ErbB2/HER2. Domain III from EGFRvIII was overlaid with (aa. 310-480) from the ErbB2/trastuzumab structure (PDB ID 1N8Z). 34E5 VHH is shown in cyan and targets a similar region of domain IV of EGFRvIII as trastuzumab for ErbB2/HER2.

REFERENCES

Adams, P. D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., & Zwart, P.H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallographica D Biol Crystallogr *66*, 213-221.

Bessman, N. J., Bagchi, A., Ferguson, K.M., & Lemmon, M.A. (2014). Complex relationship between ligand binding and dimerization in the epidermal growth factor receptor. Cell Reports *9*, 1306-1317.

Cho, H. S., Mason, K., Ramyar, K.X., Stanley, A.M., Gabelli, S.B., Denney, D.W. Jr. & Leahy, D.J. (2003). Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. Nature *421*, 756-760.

Choudhary, C., Olsen, J.V., Brandts, C., Cox, J., Reddy, P.N.G., Bohmer, F.D., Gerke, V., Schmidt-Arras, D.-E., Berdel, W.E., Muller-Tidow, C., Mann, M., & Serve, H. (2009). Mislocalized activation of oncogenic RTKs switches downstream signaling outcomes. Mol Cell *36*, 326-339.

Corbacioglu, S., Killic, M., Westhoff, M.-A., Reinhardt, D., Fulda, S., & Debatin, K.-M. (2006). Newly identified c-KIT receptor tyrosine kinase ITD in childhood AML induces ligand-independent growth and is responsive to a synergistic effect of imatinib and rapamycin. Blood *108*, 3504-3513.

Dawson, J. P., Berger, M. B., Lin, D., Schlessinger, J., Lemmon, M. A., Ferguson, K. M. (2005). EGF receptor dimerization and activation require ligand-induced conformational changes in the dimer interface. Mol Cell Biol *25*, 7734-7742.

Del Vecchio, C. A., Giacomini, C.P., Vogel, H., Jensen, K.C., Florio, T., Merlo, A., Pollack, J.R., & Wong, A.J. (2013). EGFRvIII gene rearrangement is an early event in glioblastoma tumorigenesis and expression defines a hierarchy modulated by epigenetic mechanisms. Oncogene *32*, 2670-2681.

Downward, J., Parker, P., & Waterfield, M.D. (1984). Autophosphorylation sites on the epidermal growth factor receptor. Nature *311*, 483-485.

Ekstrand, A. J., James, C.D., Cavenee, W.K., Seliger, B., Pettersson, R.F., & Collins, V.P. (1991). Genes for epidermal growth factor receptor, transforming growth factor alpha, and epidermal growth factor and their expression in human gliomas in vivo. Cancer Research *51*, 2164-2172.

Emsley, P., B. Lohkamp, W.G. Scott, and K. Cowtan (2010). Features and Development of Coot. Acta Crystallographica D *66*, 486-501.

Fan, Q. W., Cheng, C.K., Gustafson, W.C., Charron, E., Zipper, P., Wong, R.A., Chen, J., Lau, J., Knobbe-Thomsen, C., Weller, M., Jura, N., Reifenberger, G., Shokat, K.M., Weiss, W.A. (2013). EGFR phosphorylates tumor-derived EGFRvIII driving STAT3/5 and progression in glioblastoma. Cancer Cell *24*, 438-449.

Farina, A. R., Cappabianca, L., Ruggeri, P., Gneo, L., Maccarone, R., & Mackay, A.R. (2015). Retrograde TrkAIII transport from ERGIC to ER: a re-localization mechanism for oncogenic activity. Oncotarget *6*, 35636-35651.

Ferguson, K. M., Berger, M. B., Mendrola, J. M., Cho, H. S., Leahy, D. J., and Lemmon, M. A. (2003). EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. Mol Cell *11*, 507-517.

Ferguson, K. M., Darling, P. J., Mohan, M. J., Macatee, T. L., and Lemmon, M. A. (2000). Extracellular domains drive homo- but not hetero-dimerization of erbB receptors. EMBO J *19*, 4632-4643.

Gamou, S., & Shimizu, N. (1988). Glycosylation of the epidermal growth factor receptor and its relationship to membrane transport and ligand binding. J Biochem *104*, 388-396.

Gamou, S., Shimagaki, M., Minoshima, S., Kobayashi, S., & Shimizu, N. (1989). Subcellular localization of the EGF receptor maturation process. Exp Cell Res *183*, 197-206.

Gan, H. K., Cvljevic, A.N., and Johns, T.G. (2013). The epidermal growth factor receptor variant III (EGFRvIII): where wild things are altered. FEBS Journal *280*, 5350-5370.

Grandal, M. V., Zandi, R., Pedersen, M.W., Willumsen, B.M., van Deurs, B., & Poulsen, H.S. (2007). EGFRvIII escapes down-regulation due to impaired internalization and sorting to lysosomes. Carcinogeneis *28*, 1408-1417.

Greenall, S. A., Donoghue, J.F., Gottardo, N.G., Johns, T.G., Adams, T.E. (2015a). Glioma-specific domain IV EGFR cysteine mutations promote ligand-induced covalent receptor dimerization and display enhanced sensitivity to dacomitinib in vivo. Oncogene *34*, 1658-1666.

Greenall, S. A., Donoghue, J.F., Van Dineren, M., Dubljevic, V., Budiman, S., Devlin, M., Street, I., Adams, T.E., Johns, T.G. (2015b). EGFRvIII-mediated transactivation of receptor tyrosine kinases in glioma: mechanism and therapeutic implications. Oncogene *34*, 5277-5287.

Heimberger, A. B., Hlatky, R., Suki, D., Yang, D., Weinberg, J., Glblert, M., Sawaya, R., & ALdape, K. (2005). Prognostic effect of epidermal growth factor receptor and EGFRvIII in glioblastoma multiforme patients. Clinical Cancer Research *11*, 1462-1466.

Huang, P. H., Mukasa, A., Bonavia, R., Flynn, R.A., Brewer, Z.E., Cavenee, W.K., Furnari, F.B., White, F.M. (2007). Quantitative analysis of EGFRvIII cellular signaling networks reveals a combinatorial therapeutic strategy for glioblastoma. Proc Natl Acad Sci USA *104*, 12867-12872.

Johns, T. G., Stockert, E., Ritter, G., Jungbluth, A.A., Huang, H.J., Cavenee, W.K., Smyth, F.E., Hall, C.M., Watson, N., Nice, E.C., Gullick, W.J., Old, L.J., Burgess, A.W., & Scott, A.M. (2002). Novel monoclonal antibody specific for the del2-7 epidermal growth factor receptor (EGFR) that also recognizes the EGFR expressed in cells containing amplification of the EGFR gene. Int J Cancer *98*, 398-408.

Johnson, H., Del Rosario, A.M., Bryson, B.D., Schroeder, M.A., Sarkaria, J.N., & White, F.M. (2012). Molecular Characterization of EGFR and EGFRvIII Signaling Networks in Human Glioblastoma Xenografts. Mol Cell Proteomics *11*, 1724-1740.

Johnson, L. A., Scholler, J., Ohkuri, T., Kosaka, A., Patel, P.R., McGettigan, S.E., Nace, A.K., Dentchev, T., Thekkat, P., Loew, A., Boesteanu, A.C., Cogdill, A.P., Chen, T., Fraietta, J.A., Kloss, C.C., Posey, A.D., Jr., Engels, B., Singh, R., Ezell, T., Idamakanti, N., Ramones, M.H., Li, N., Zhou, L., Plesa, G., Seykora, J.T., Okada, H., June, C.H., Brogdon, J.L., & Maus, M.V. (2015). Rational development and characterization of humanized anti-EGFR variant III chimeric antigen receptor T cells for glioblastoma. Sci Transl Med *7*, 275ra222.

Kaszuba, K., Grzybek, M., Orlowski, A., Danne, R., Róg, T., Simons, K., Coskun, U., & Vattulainen, I. (2015). N-Glycosylation as determinant of epidermal growth factor receptor conformation in membranes. Proc Natl Acad Sci USA *112*, 4334-4339.

Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. J Appl Crystallogr *36*, 1277-1282.

Lawrence, M. C., Colman, P.M. (1993). Shape complementarity at protein/protein interfaces. J Mol Biol *234*, 946-950.

Lee, S., Greenlee, E.B., Amick, J.R., Ligon, G.F., Lillquist, J.S., Natoli, E.J. Jr., Hadari, Y., Alvarado, D., Schlessinger, J. (2015). Inhibition of ErbB3 by a monoclonal antibody that locks the extracellular domain in an inactive configuration. Proc Natl Acad Sci USA *112*, 13225-13230.

Lemmon, M. A., Schlessinger, J., & Ferguson, K.M. (2014). The EGFR family: not so prototypical receptor tyrosine kinases. Cold Spring Harb Perspect Biol *6*, a020768.

Li, S., Kussie, P., Ferguson, K.M. (2008). Structural basis for EGF receptor inhibition by the therapeutic antibody IMC-11F8. Structure *16*, 216-227, doi: 210.1016/j.str.2007.1011.1009.

Li, S., Schmitz, K. R., Jeffrey, P. D., Wiltzius, J. J., Kussie, P., and Ferguson, K. M. (2005). Structural basis for inhibition of the epidermal growth factor receptor by cetuximab. Cancer Cell *7*, 301-311.

Li, S., Schmitz, K.R., Jeffrey, P.D., Wiltzius, J.J., Kussie, P., Ferguson, K.M. (2005). Structural basis for inhibition of the epidermal growth factor receptor by cetuximab. Cancer Cell *7*, 301-311.

Libermann, T. A., Nusbaum, H.R., Razon, N., Kris, R., LAx, I., Soreq, H., Whittle, N., Waterfield, M.D., Ulrich, A., & Schlessinger, J. (1985). Amplification, enhanced expression and possible rearrangement of EGF receptor gene in primary human brain tumours of glial origin. Nature *313*, 144-147.

Liu, Y.-C., Yen, H.-Y., Chen, C.-Y., Chen, C.-H., Cheng, P.-F., Juan, Y.-H., Chen, C.-H., Khoo, K.-H., Yu, C.-J., Yang, P.-C., Hsu, T.-L., & Wong, C.-H. (2011). Sialylation and fucosylation of epidermal growth factor receptor suppresses its dimerization and activation in lung cancer cells. Proc Natl Acad Sci USA *108*, 11332-11337.

Luwor, R. B., Johns, T.G., Murone, C., Huang, H.J., Cavenee, W.K., Ritter, G., Old, L.J., Burgess, A.W., & Scott, A.M. (2001). Monoclonal antibody 806 inhibits the growth of tumor xenografts expressing either the del2-7 or amplified epidermal growth factor receptro (EGFR) but not wild-type EGFR. Cancer Research *61*, 5355-5361.

Maley, F., Trimble, R.B., Tarentino, A.L. Plummer, T.H., Jr. (1989). Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases. Anal Biochem *180*, 195-204.

McCoy, A. J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J Appl Cryst *40*, 658-674.

Morin, A., Eisenbraun, B., Key, J., Sanschagrin, P.C., Timony, M.A., Ottaviano, M., & Sliz, P. (2013). Collaboration gets the most out of software. Elife *2*.

Nag, M., Bera, K., & Basak, S. (2015). Intermolecualr disulide bond formation promotes immunoglobulin aggregation: investigation by fluorescence correlation spectroscopy. Proteins *83*, 169-177.

O'Rourke, D. M., Nute, E.J., Davis, J.G., Wu, C., Lee, A., Murali, R., Zhang, H.T., QIan, X., Kao, C.C., & Greene, M.I. (1998). Inhibition of a naturally occurring EGFR oncoprotein by the p185neu ectodomain: implications for subdomain contributions to receptor assembly. Oncogene *16*, 1197-1207.

Oslowski, C. M., & Urano, F. (2011). Measuring ER stress and the unfolded protein response using mammalian tissue culture system. Methods Enzymol *490*, 71-92.

Pillay, V., Allaf, L., Wilding, A.L., Donoghue, J.F., Court, N.W., Greenall, S.A., et al. (2009). The plasticity of oncogene addiction: implications for targeted therapies directed to receptor tyrosine kinases. Neoplasia *11*, 442-458.

Reilly, E. B., Phillips, A.C., Buchanan, F.G., Kingsbury, G., Zhang, Y., Meulbroek, J.A., Cole, T.B., DeVries, P.J., Falls, H.D., Beam, C., Gu, J., Digiammarino, E.L., Palma, J.P., Donawho, C.K., Goodwin, N.C., & Scott, A.M. (2015). Characterization of ABT-806, a humanized tumor-specific anti-EGFR monoclonal antibody. Mol Cancer Ther *14*, 1141-1151.

Schmidt-Arras, D., Bohmer, S.-A., Koch, S., Muller, J.P., Blei, L., Cornils, H., Bauer, R., Korasikha, S., Thiede, C., & Bohmer, F.-D. (2009). Anchoring of FLT3 in the endoplasmic reticulum alters signaling quality. Blood *113*, 3568-3576.

Schmitz, K. R., & Ferguson, K.M. (2009). Interaction of antibodies with ErbB receptor extracellular regions. Exp Cell Res *315*, 659-670.

Schmitz, K. R., Bagchi, A., Roovers, R.C., van Bergen en Henegouwen, P.M., & Ferguson, K.M. (2013). Structural evaluation of EGFR inhibition mechanisms for nanobodies/VHH domains. Structure *21*, 1214-1224.

Shinojima, N., Tada, K., Shiraishi, S., Kamiryo, T., Kochi, M., Nakamura, H., Makino, K., Saya, H., Hirano, H., Kuratsu, J.-I., et al. (2003). Prognostic value of epidermal growth factor receptor in patients with glioblastoma multiforme. Cancer Research *63*, 6962-6970.

Studier, F. W. (2005). Protein production by auto-induction in high density shaking cultures. Protein Expr Purif *41*, 207-234.

Sugawa, N., Ekstrand, A. J., James, C. D., and Collins, V. P. (1990). Identical splicing of aberrant epidermal growth factor receptor transcripts from amplified rearranged genes in human glioblastomas. Proc Natl Acad Sci USA *87*, 8602-8606.

Svergun, D. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J Appl Crystallogr *25*, 495-503.

Svergun, D. I., Barberato, C., & Koch, M.H.J. (1995). CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. J Appl Cryst *28*, 768-773.

Takasu, S., Takahashi, T., Okamoto, S., Oriuchi, N., Nakayashiki, N., Okamoto, K., Muramatsu, H., Hayashi, T., Nakahara, N., Mizuno, M., Wakabayashi, T., Higuchi, T., Endo, K., Kozaki, K., Miyaishi, O., Saga, S., Ueda, R., Yoshida, J. & Yoshikawa, K. (2003). Radioimmunoscintigraphy of intracranial glioma xenograft with a technetium-99m-labeled mouse monoclonal antibody specifically recognizing type III mutant epidermal growth factor receptor. J Neurooncol *63*, 247-256.

Tarentino, A. L. P., T.H., Jr., & Maley, F. (1974). The release of intact oligosaccharides from specific glycoproteins by endo-beta-N-acetylglucosaminidase H. J Biol Chem *249*, 818-824.

Terwilliger, T. C., Grosse-Kunstleve, R.W., Afonine, P.V., Moriarty, N.W., Zwart, P.H., Hung, L.W., Read, R.J., & Adams, P.D. (2008). Iterative model building, structure refinement, and density modification with the PHENIX AutoBuild wizard. Acta Crystallographica D Biol Crystallogr *64*, 61-69.

Tsuda, T., Ikeda, Y., Taniguchi, N. (2000). The Asn-420 linked sugar chain in human epidermal growth factor receptor suppresses ligand-independent spontaneous oligomerization. Possible role of a specific sugar chain in controllable receptor activation. J Biol Chem *275*, 21988-21994.

Wikstrand, C. J., Hale, L.P., Batra, S.K., Hill, M.L., Humphrey, P.A., Kurpad, S.N., et all. (1995). Monoclonal antibodies against EGFRvIII are tumor specific and react with breast and lung carcinomas and malignant gliomas. Cancer Research *55*, 3140-3148.

Wikstrand, C. J., McLendon, R.E., Friedman, A.H., & Bigner, D.D. (1997). Cell surface localization and density of the tumor-assocaited variant of the epidermal growth factor receptor, EGFRvIII. Cancer Research *57*, 4130-4140.

Wong, A. J., Ruppert, J.M., Bigner, S.H., Grzeschik, C.H., Humphrey, P.A., Bigner, D.S., & Vogelstein, B. (1992). Structural alterations of the epidermal growth factor receptor gene in human gliomas. Proc Natl Acad Sci USA *89*, 2965-2969.

Ymer, S. I., Greenall, S.A., Cvrljevic, A., Cao, D.X., Donoghue, J.F., Epa, V.C., Scott, A.M., Adams, T.E., & Johns, T.G. (2011). Glioma Specific Extracellular Missense Mutations in the First Cysteine Rich Region of Epidermal Growth Factor Receptor (EGFR) Initiate Ligand Independent Activation. Cancers (Basel) *3*, 2032-2049.

CHAPTER SIX

Conclusions and Perspectives

In this dissertation, I have presented work that identifies novel extracellular mechanisms of oncogenic dysregulation of the Epidermal Growth Factor receptor (EGFR). One such extracellular mechanism, as discussed in Chapter Two, is enhancement of ligand binding affinity without measureable alteration of homo-dimerization energetics *in vitro* (Bessman, 2014). The linkage of ligand binding and dimerization equilibria in the EGFR has previously been thought to be weak (Johannessen, 2001). We report that it is competitive, and that these extracellular GBM mutations may alter the thermodynamic linkage of ligand binding and receptor dimerization (Bessman, 2014). These observations can be structurally rationalized by the view that domain I/II interactions function normally to restrict the conformation of domain I, and breaking these interactions make it energetically more favorable for domain I to come closer to domain III to form a high affinity ligand binding site. It is possible that alterations in the linkage of the thermodynamic equilibria may alter relationships between homo- and hetero-oligomerization energetics of ErbB/HER receptors in the context of the intact full-length receptors expressed on the cell surface, thereby stabilizing distinct receptor dimer species that may contribute to constitutive signaling. Stabilization of these alternate heterodimeric species may mimic the effects of some EGFR ligands, such as EGF, which binds preferentially to EGFR/ErbB2 heterodimers (Macdonald-Obermann, 2014).

Another possibility is that contributions from the membrane and/or the intracellular region of EGFR are key to this oncogenic mechanism, and is lacking in our current analyses. Such non-additivity or interdependence of contributions from the extra- and intra-cellular regions of EGFR would be consistent with the observation that negative cooperativity of intact EGFR is lost in studies of the isolated extracellular region of human EGFR, but not in the *Drosophila melanogaster* ortholog (Alvarado et al., 2010). Such allosteric linkage is also implicit in the observation that these extracellular activating mutations alter the differential sensitivity of EGFR inhibition by lapatinib and erlotinib. Lapatinib, binds preferentially to the

kinase inactive conformation and has slow off-rate relative to gefitinib (Wood, 2004), where erlotinib can bind both active and inactive kinase conformations (Park, 2012). These extracellular, activating mutations therefore sensitize EGF receptor to an inhibitor known to bind preferentially to the kinase inactive conformation, suggesting that structural allostery is at play in this extracellular mode of oncogenic dysregulation. Experimental tools for evaluating allostery and the relative contributions of extra- and intra-cellular interactions to receptor-mediated dimerization are inherently limited by the size of the receptor and the quantities that can be purified in a cost-effective manner. In Chapter Three, I have presented preliminary progress towards application of hydrogen/deuterium exchange coupled to mass spectrometry towards understanding the relative contributions of extra- and intra-cellular interactions to dimerization energetics. It is possible that extracellular activating mutations alter the linkage of extra- and intra-cellular dimerization energy rather than just simply enhancing extracellular-mediated dimerization, and HDX/MS would be a useful tool to study this mode of dysregulation. Specifically, HDX/MS analyses of ligand-induced protection of full-length EGFR could directly report on dimerization energetics in peptides that are known from crystallographic analyses to be in either extra- or intra-cellular dimerization interfaces. Analyses of receptor dynamics through both experimental and computational techniques has recently progressed our understanding of transmembrane regulation of the EGFR (Arkhipov et al., 2013; Endres et al., 2013), as well as mechanisms by which it can be oncogenically dysregulated by mutations in its tyrosine kinase domain (Shan, 2012).

Another extracellular oncogenic mechanism, as discussed in Chapter Five, is an alteration of trafficking and increased flexibility of domain II and altered inter-domain relationships. Deletion of parts of the extracellular region of EGFR can cause improper folding of the intact receptor, thereby causing its sequestration of its proper trafficking to the plasma membrane, as has previously been observed for EGFRvIII (Wikstrand, 1997). In previous work, the Ferguson laboratory in collaboration in the Carpenter laboratory, showed

that the unpaired cysteine of EGFRvIII causes a similar sequestration of these mutated receptor molecules early in the secretory pathway through disulfide-mediate dimerization, consistent with recently published work (Ymer, 2011). I compliment these data with solution scattering analysis of the EGFRvIII extracellular region, which suggest intrinsic structural flexibility in domain II, consistent with its low X-ray scattering properties in two X-ray structures. This alteration of domain II conformation and dynamics may lower the energetic barrier associated with hetero-dimerization, rationalizing published reports that EGFRvIII hetero-dimerizes with and trans-activates other receptor tyrosine kinase, such as MET (Huang, 2007; Johnson, 2012). This transactivation relies on activity of EGFRvIII, which is not understood in structural terms. However, a separate report suggests that the constitutive activation of EGFRvIII relies on co-expression and activity of wild-type EGFR (Fan, 2013). Taken together, both observations would be consistent with the view that EGFRvIII dynamically hetero-oligomerizes with other receptor tyrosine kinases, including both EGFR and MET (Greenall, 2015). The precise structural basis for this hetero-oligomerization, and whether EGFRvIII is a preferred 'receiver' or preferred 'activator' in the asymmetric dimer (Zhang, 2006), and the mechanism of cooperation between EGFRvIII and EGFR, remains unknown. It is clear, however, that this hetero-oligomerization requires activity of EGFRvIII (Greenall, 2015), in contrast to ErbB2-hetero-olgiomerization with EGF receptor. Specifically, kinase activity of ErbB2 is not required for its hetero-oligomerization with and modulation of EGFR activity (Qian, 1994a; Qian, 1994b). There also remain formidable technical challenges to being able to address the question of how EGFRvIII transactivates and cooperates with other RTK's, as a result of the range of structural techniques that can currently be used on intact membrane proteins of a similar size as EGFRvIII.

Both oncogenic mechanisms rely on alteration of structural relationships between the subdomains of the EGFR extracellular region. The point mutations studied in Chapter Two disrupt autoinhibitory interactions between domains I and II that likely serve to restrict the

conformation of domain II in an inactive state in the inactive 'tethered' structure. Unrestraining the domain I/II relationship therefore may allow domain II to adopt conformations distinct from those observed in 'tethered' X-ray crystal structures. Loss of these autoinhibitory interactions therefore may lower the energetic barrier for domain II conformational remodeling, that is required for ErbB homo-dimerization or hetero-dimerization (Dawson, 2005; Dawson et al., 2007) and consequently result in constitutive receptor activation and oncogenic signaling.

In a second focus, I studied two distinct mechanisms of antibody recognition of the EGFR – one in the context of resistance to a currently clinically used antibody drug, and the other in the context of an oncogenic alternative splice variant that renders a unique, tumor-specific antigen. In the first focus, my work on the structural basis for high affinity necitumumab binding to the S468R cetuximab resistance mutation identifies a hydrophobic cavity in between the VH and VL domains in necitumumab that is a result of CDR H3 conformation. This hydrophobic cavity can accommodate genetic mutations that introduce bulky residues, where cetuximab cannot as a result of its flatter paratope. In my X-ray crystal structure of isolated EGFR domain III containing the S468R mutation bound the Fab fragment from necitumumab (Fab11F8), this cavity contains a buried water molecule that is coordinated by a serine in the VH domain that, along with several cation-pi and polar interactions, stabilizes the arginine sidechain in this region of low dielectric constant. Analysis of the four complexes in the asymmetric unit identified a structural 'wobble', suggesting that the antibody is less rigidly bound to the mutated EGF receptor domain III compared to the wild-type domain III (Li, 2008).

Although the presence of this buried hydrophobic cavity does not alter the common quantitative measure for shape complementarity of protein/protein interfaces (Lawrence, 1993), comparison of the paratopes of cetuximab and necitumumab to those of other therapeutic antibodies reveals a structural class of antibodies that contain similar buried

cavities, as opposed to the 'flatter' paratope of some antibodies such as cetuximab. The origin of this cavity is due primarily to CDR H3 conformation, which is known to be hypervariable and adopt a structurally diverse set of conformations, driven by a C-terminal kink (Sircar, 2009; Weitzner, 2015). It is also thought that affinity maturation of antibodies can 'bias' the conformation of CDR H3 (Xu, 2015). These studies highlight the importance of development of computational tools to classify CDR H3 conformation and distinguish surface features of antibody paratopes. Some existing tools, such as contact analysis (MacCallum, 1996), may be useful for identifying whether newer antibodies contain similar cavities. Further studies are needed to evaluate whether antibodies that contain similar cavities are capable of exhibiting a similar degree of structural plasticity as necitumumab, and whether they are less susceptible to resistance through epitope mutations. Our analysis also reveals that paratope shape may also be a property of antibodies that could be utilized in synergistic modes for combinatorial therapies.

I also discovered a mechanism by which antibody specificity for a commonly found oncogenic EGFR variant, EGFRvIII, can be engineered. EGFRvIII is a target antigen for many evolving cancer immunotherapies, including both a CD3-bispecific antibody (Choi, 2013) as well as chimeric antigen receptor expressing T cells (Johnson, 2015). I discovered in an X-ray crystal structure of an EGFRvIII-specific heavy chain only antibody or VHH domain that the VHH domain functionally mimics the intramolecular 'tethering' interaction present in wild-type EGFR but absent in EGFRvIII as a result of its exclusion of mature amino acids 6-273. The VHH targets an epitope on EGFR domain IV that is sterically occluded as a result of this intramolecular 'tethering' interaction, thereby gaining specificity for EGFRvIII. This EGFRvIII-targeted VHH domain therefore gains specificity for EGFRvIII in a manner analogous to how trastuzumab/Herceptin™ and pertuzumab/Perjeta™ gain specificity for ErbB2/HER2, as a result of it being constitutively 'extended' as a monomer (Cho, 2003; Franklin, 2004). Trastuzumab targets a very analogous epitope on ErbB2 domain IV, and

pertuzumab targets an epitope on domain II, both of which are sterically occluded in EGFR, ErbB3, and ErbB4 as a result of the presence of the intramolecular 'tether' between domains II and IV in the absence of ligand (Bouyain, 2005; Cho, 2002; Ferguson et al., 2003).

This EGFRvIII-specific VHH domain may be useful as an imaging agent, or in therapeutic applications. VHH domains, as a result of their small size, have improved tumor and tissue penetration properties compared to larger monoclonal antibodies, and are emerging as potential imaging agents (Schmitz, 2013). As VHH domains can also be used as 'cassettes' or modules to combine synergistic paratopes into one agent (Roovers, 2011; Schmitz, 2013), specific recognition of EGFRvIII with the 34E5 VHH domain could potentially be combined with other paratopes that recognize tumor-specific cell surface antigens, or in the context of chimeric antigen receptor expressing T cells. Identification of this novel mechanism by which oncogenic EGFR variant III can be specifically targeted is important for the development of this VHH domain for these potential applications.

Overall, this dissertation has contributed to our understanding of both extracellular mechanisms and antibody targeting of the epidermal growth factor receptor family. The work in the first part leads us towards being able to ask questions about how the epidermal growth factor receptor functions as an intact molecule. The first step in understanding allosteric regulation of the ErbB/HER family is to understand the allosteric alterations of the intact receptor imposed by ligand binding within an EGFR homodimer. Development of structural techniques to analyze transmembrane allosteric regulation of the ErbB/HER family of receptor tyrosine kinases (RTK's) are ongoing, and hydrogen/deuterium exchange coupled to mass spectrometry holds promise for addressing how extracellular oncogenic mechanisms relay information across the plasma membrane. Future studies that may build upon the work presented in this dissertation may rationalize differences in tyrosine kinase inhibitor sensitivity and inform future development of antibodies or other therapeutic agents with engineered specificity for these mutated receptor molecules.

Our preliminary advances in understanding extracellular mechanisms by which the Epidermal Growth Factor receptor can be allosterically dysregulated may have lessons that are general for how transmembrane proteins are allosterically regulated in general. An alteration of coupling of thermodynamic equilibria may stabilize distinct heteromeric species on the cell surface and alter trafficking of these receptors, for example, and thereby cause oncogenic signaling. Our efforts at understanding transmembrane alllostery by exploiting hydrogen/deuterium exchange coupled to mass spectrometry could impact future efforts at understanding structural allostery in membrane proteins using HDX/MS. In principle, membrane bound receptors are in many ways similar to allosterically regulated enzymes for which mechanisms have been described over the past half century (Levitzki, 1969). The distinction between such enzymes and receptor tyrosine kinases, however, is the presence in the latter of a transmembrane domain as well as a membrane composed of an asymmetric lipid composition through which such allosteric information may be relayed. Juxtamembrane regions of RTK's, which may interact directly with these lipid headgroups, and thereby play roles in amplifying these signals— leading to structural arrangements that favor autoinhibited or activated states of these receptors. Indeed, in the EGFR, many interactions between the juxtamembrane regions and kinase (Jura, 2009; Red Brewer, 2009), and Thr654 phosphorylation are already known to affect transmembrane communication (Defize, 1989; Felder, 1992; Lund, 1990). For example, Protein Kinase C phosphorylation of Thr654 inhibits high affinity EGF binding and slows EGFR internalization (Lund, 1990). These juxtamembrane interactions may serve to restrict the conformational freedom of the tyrosine kinase domains until they are relieved by ligand binding. There are other receptor systems that are thought to be allosterically regulated for which interactions with the membrane may also be autoinhibitory, such as the T-cell receptor, in which Wucherpfennig and colleagues have proposed that the cytoplasmic tyrosine-based motifs (ITAM's) are buried in the membrane prior to release by peptide-MHC engagement (Xu, 2008).

Another possible mode of allosteric dysregulation is ‘unlinkage’ of the relative contributions of extracellular and intracellular contributions to receptor dimerization. Such non-additivity of extra- and intra-cellular energetics within receptor dimers harkens back to nonadditivity principles in closed thermodynamic systems (Mark, 1994), or context-dependent effects of a thermodynamic perturbation on a closed system, which in this case is ligand binding. Preliminary progress presented in this dissertation could be expanded to utilize HDX/MS to measure the simultaneous ligand-induced protection factors in peptides that are known to be in dimerization interfaces. Such relative protection factors could be interpreted as relative contributions of extra- and intra-cellular dimerization energetics, and used to compare how a mutated receptor would ‘spread’ its dimerization energy relative to a wild-type receptor.

The second part of this dissertation identifies new mechanisms of plasticity and specificity for antibody targeting of the EGFR, either in the context of retaining high affinity for mutantions that cause resistance to currently used therapeutic antibodies like cetuximab, or in the context of specifically targeting oncogenic variants of the EGFR, like EGFRvIII. Our work in the first part both identified a structural mechanism for necitumumab inhibition of EGF receptor variants that cause acquired resistance to cetuximab. Our analysis suggests that necitumumab belongs to a structural class of antibodies that may be defined by conformation of CDR H3 that contain buried cavities between their VH and VL domains. These cavities facilitate structural plasticity in accommodating genetic mutations that may cause a decrease in binding energetics to an antibody with a relative flat paratope incapable of accommodating such changes. Therefore, our analysis suggests that paratope shape may be a useful property of antibodies that could be used in synergistic modes in the context of resistance to one therapeutic antibody. My second focus corroborates the utility of the extracellular ‘tether’ interaction in the ErbB/HER family for gaining antibody specificity to an oncogenic EGFR variant. Both focuses reveal that antibody specificity and plasticity are not mutually exclusive,

and future work will have to be done on these and other antibody systems to reveal the limits of how much affinity or specificity can be achieved without compromising plasticity.

REFERENCES

Alvarado, D., Klein, D. E., and Lemmon, M. A. (2010). Structural Basis for Negative Cooperativity in Growth Factor Binding to an EGF Receptor. Cell *142*, 568-579.

Arkhipov, A., Shan, Y., Das, R., Endres, N. F., Eastwood, M. P., Wemmer, D. E., Kuriyan, J., and Shaw, D. E. (2013). Architecture and membrane interactions of the EGF receptor. Cell *152*, 557-569.

Bessman, N. J., Bagchi, A., Ferguson, K.M., & Lemmon, M.A. (2014). Complex relationship between ligand binding and dimerization in the epidermal growth factor receptor. Cell Reports *9*, 1306-1317.

Bouyain, S., Longo, P.A., Li, S., Ferguson, K.M., Leahy, D.J. (2005). The extracellular region of ErbB4 adopts a tethered conformation in the absence of ligand. PNAS *102*, 15024-15029.

Cho, H. S., & Leahy, D.J. (2002). Structure of the extracellular region of HER3 reveals an intermolecular tether. Science *297*, 1330-1333.

Cho, H. S., Mason, K., Ramyar, K.X., Stanley, A.M., Gabelli, S.B., Denney, D.W. Jr. & Leahy, D.J. (2003). Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. Nature *421*, 756-760.

Choi, B. D., Kuan, C.-T., Cai, M., Archer, G.E., Mitchell, D.A., Gedeon, P.C., Sanchez-Perez, L., Pastan, I., Bigner, D.D., & Sampson, J.H. (2013). Systemic administration of a bispecific antibody targeting EGFRvIII successfully treats intracerebral glioma. Proc Natl Acad Sci USA *110*, 270-275.

Dawson, J. P., Berger, M. B., Lin, D., Schlessinger, J., Lemmon, M. A., Ferguson, K. M. (2005). EGF receptor dimerization and activation require ligand-induced conformational changes in the dimer interface. Mol Cell Biol *25*, 7734-7742.

Dawson, J. P., Bu, Z., and Lemmon, M. A. (2007). Ligand-Induced Structural Transitions in ErbB Receptor Extracellular Domains. Structure *15*, 942-954.

Defize, L. H., Boonstra, J., Meisenhelder, J., Kruijer, W., Tertoolen, L.G., Tilly, B.C., Hunter, T., van Bergen en Henegouwen, P.M., Moolenaar, W.H., & de Laat, S.W. (1989). Signal transduction by epidermal growth factor occurs through the subclass of high affinity receptors. J Cell Biol *109*, 2495-2507.

Endres, N. F., Das, R., Smith, A. W., Arkhipov, A., Kovacs, E., Huang, Y., Pelton, J. G., Shan, Y., Shaw, D. E., Wemmer, D. E.*, et al.* (2013). Conformational coupling across the plasma membrane in activation of the EGF receptor. Cell *152*, 543-556.

Fan, Q. W., Cheng, C.K., Gustafson, W.C., Charron, E., Zipper, P., Wong, R.A., Chen, J., Lau, J., Knobbe-Thomsen, C., Weller, M., Jura, N., Reifenberger, G., Shokat, K.M., Weiss, W.A. (2013). EGFR phosphorylates tumor-derived EGFRvIII driving STAT3/5 and progression in glioblastoma. Cancer Cell *24*, 438-449.

Felder, S., LaVin, J., Ullrich, A., & Schlessinger, J. (1992). Kinetics of binding, endocytosis, and recycling of EGF receptor mutants. J Cell Biol *117*, 203-212.

Ferguson, K. M., Berger, M. B., Mendrola, J. M., Cho, H. S., Leahy, D. J., and Lemmon, M. A. (2003). EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. Mol Cell *11*, 507-517.

Franklin, M. C., Carey, K.D., Vajdos, F.F., Leahy, D.J., de Vos, A.M., & Sliwkowski, M.X. (2004). Insights into ErbB signaling from the structure of the ErbB2-pertuzumab complex. Cancer Cell *5*, 317-328.

Greenall, S. A., Donoghue, J.F., Van Dineren, M., Dubljevic, V., Budiman, S., Devlin, M., Street, I., Adams, T.E., Johns, T.G. (2015). EGFRvIII-mediated transactivation of receptor tyrosine kinases in glioma: mechanism and therapeutic implications. Oncogene *34*, 5277-5287.

Huang, P. H., Mukasa, A., Bonavia, R., Flynn, R.A., Brewer, Z.E., Cavenee, W.K., Furnari, F.B., White, F.M. (2007). Quantitative analysis of EGFRvIII cellular signaling networks reveals a combinatorial therapeutic strategy for glioblastoma. Proc Natl Acad Sci USA *104*, 12867-12872.

Johannessen, L. E., Haugen, K.E., ostvold, A.C., Stang, E., & Madshus, I.H. (2001). Heterodimerization of the epidermal growth factor (EGF) receptor and ErbB2 and the affinity of EGF binding are regulated by different mechanisms. Biochemical Journal *356*, 87-96.

Johnson, H., Del Rosario, A.M., Bryson, B.D., Schroeder, M.A., Sarkaria, J.N., & White, F.M. (2012). Molecular Characterization of EGFR and EGFRvIII Signaling Networks in Human Glioblastoma Xenografts. Mol Cell Proteomics *11*, 1724-1740.

Johnson, L. A., Scholler, J., Ohkuri, T., Kosaka, A., Patel, P.R., McGettigan, S.E., Nace, A.K., Dentchev, T., Thekkat, P., Loew, A., Boesteanu, A.C., Cogdill, A.P., Chen, T., Fraietta, J.A., Kloss, C.C., Posey, A.D., Jr., Engels, B., Singh, R., Ezell, T., Idamakanti, N., Ramones, M.H., Li, N., Zhou, L., Plesa, G., Seykora, J.T., Okada, H., June, C.H., Brogdon, J.L., & Maus, M.V. (2015). Rational development and characterization of humanized anti-EGFR variant III chimeric antigen receptor T cells for glioblastoma. Sci Transl Med *7*, 275ra222.

Jura, N., Endres, N.F., Engel, K., Deindl, S., Das, R., Lamers, M.H., Wemmer, D.E., Zhang, X., Kuriyan, J. (2009). Mechanism for activation of the EGF receptor catalytic domain by the juxtamembrane segment. Cell *137*, 1293-1307.

Lawrence, M. C., Colman, P.M. (1993). Shape complementarity at protein/protein interfaces. J Mol Biol *234*, 946-950.

Levitzki, A., Koshland, D.E., Jr. (1969). Negative cooperativity in regulatory enzymes. Proc Natl Acad Sci USA *62*, 1121-1128.

Li, S., Kussie, P., Ferguson, K.M. (2008). Structural basis for EGF receptor inhibition by the therapeutic antibody IMC-11F8. Structure *16*, 216-227, doi: 210.1016/j.str.2007.1011.1009.

Lund, K. A., Lazar, C.S., Chen, W.S., Walsh, B.J., Welsh, J.B., Herbst, J.J., Walton, G.M., Rosenfeld, M.G., Gill, G.N., & Wiley, H.S. (1990). Phosphorylation of the epidermal growth

factor receptor at threonine 654 inhibits ligand-induced internalization and down-regulation. J Biol Chem *265*, 20517-20523.

MacCallum, R. M., Martin, A.C., & Thornton, J.M. (1996). Antibody-antigen interactions: contact analysis and binding site topography. J Mol Biol *262*, 732-745.

Macdonald-Obermann, J. L., & Pike, L.J. (2014). Different epidermal growth factor (EGF) receptor ligands show distinct kinetics and biased or partial agonism for homodimer and heterodimer. J Biol Chem *289*, 26178-26188.

Mark, A. E., & van Gunsteren, W.F. (1994). Decomposition of the Free Energy of a System in Terms of Specific Interactions: Implications for Theoretical and Experimental Studies. J Mol Biol *240*, 167-176.

Park, J. H., Liu, Y., Lemmon, M.A., & Radhakrishnan, R. (2012). Erlotinib binds both inactive and active conformations of the EGFR tyrosine kinase domain. Biochemical Journal *448*, 417-423.

Qian, X., Dougall, W.C., Hellman, M.E. & Greene, M.I. (1994a). Kinase-deficient neu proteins suppress epidermal growth factor receptor function and abolish cell transformation. Oncogene *9*, 1507-1514.

Qian, X., LeVea, C.M., Freeman, J.K., Dougall, W.C., & Greene, M.I. (1994b). Heterodimerization of epidermal growth factor receptor and wild-type or kinase-deficient Neu: a mechanism of interreceptor kinase activation and transphosphorylation. Proc Natl Acad Sci USA *91*, 1500-1504.

Red Brewer, M., Choi, S.H., Alvarado, D., Moravcevic, K., Pozzi, A., Lemmon, M.A., & Carpenter, G. (2009). The juxtamembrane region of the EGF receptor functions as an activation domain. Mol Cell *34*, 641-651.

Roovers, R. C., Vosjan, M.J., Laeremans, T., el Khoulati, R., de Bruin, R.C., Ferguson, K.M., Verkleij, A.J., van Dongen, G.A., & van Bergen en Henegouwen, P.M. (2011). A biparatopic anti-EGFR nanobody efficientyl inhibits solid tumour growth. Int J Cancer *129*, 2013-2024.

Schmitz, K. R., Bagchi, A., Roovers, R.C., van Bergen en Henegouwen, P.M., & Ferguson, K.M. (2013). Structural evaluation of EGFR inhibition mechanisms for nanobodies/VHH domains. Structure *21*, 1214-1224.

Shan, Y., Eastwood, M.P., Zhang, X., Kim, E.T., Arkhipov, A., Dror, R.O., Jumper, J., Kuriyan, J., & Shaw, D.E. (2012). Oncogenic mutations counteract intrinsic disorder in the EGFR kinase and promote receptor dimerization. Cell *149*, 860-870.

Sircar, A., Kim, E.T., & Gray, J.J. (2009). RosettaAntibody: antibody variable region homology modeling server. Nucleic Acids Res *37*, W474-W479.

Weitzner, B. D., Dunbrack, R.L., Jr., & Gray, J.J. (2015). The origin of CDR H3 structural diversity. Structure *23*, 302-311.

Wikstrand, C. J., McLendon, R.E., Friedman, A.H., & Bigner, D.D. (1997). Cell surface localization and density of the tumor-assocaited variant of the epidermal growth factor receptor, EGFRvIII. Cancer Research *57*, 4130-4140.

Wood, E. R., Truesdale, A.T., McDonald, O.B., Yuan, D., Hassell, A., Dickerson, S.H., Ellis, B., Pennisi, C., Home, E., Lackey, K., Alligood, K.J., Rusnak, D.W., Gilmer, T.M., & Shewchuk, L. (2004). A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells. Cancer Research *64*, 6652-6659.

Xu, C., Gagnon, E., Call, M.E., Schnell, J.R., Schieters, C.D., Carman, C.V., Chou, J.J. & Wucherpfennig, K.W. (2008). Regulatipn of T cell receptor activation by dynamic membrane binding of the CD3epsilon cytoplasmic tyrosine-based motif. Cell *135*, 702-713.

Xu, H., Schmidt, A.G., O'Donnell, T., Therkelsen, M.D., Kepler, T.B., Moody, M.A., Haynes, B.F., Liao, H.X., Harrison, S.C., & Shaw, D.E. (2015). Key mutations stabilize antigen-binding conformation during affinity maturation of a broadly neutralizing influenza antibody lineage. Proteins *83*, 771-780.

Ymer, S. I., Greenall, S.A., Cvrljevic, A., Cao, D.X., Donoghue, J.F., Epa, V.C., Scott, A.M., Adams, T.E., & Johns, T.G. (2011). Glioma Specific Extracellular Missense Mutations in the First Cysteine Rich Region of Epidermal Growth Factor Receptor (EGFR) Initiate Ligand Independent Activation. Cancers (Basel) *3*, 2032-2049.

Zhang, X., Gureasko, J., Shen, K., Cole, P.A., Kuriyan, J. (2006). An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. Cell *125*, 1137-1149.